universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Volatility Derivatives: Valuation and Hedging

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag.rer.nat.)

Verfasser:	Michael Punz
Matrikel-Nummer:	A 0648834
Studienrichtung:	A 405 Mathematik
Betreuer:	Univ. Prof. Dr. Walter Schachermayer
	Ass. Prof. Dr. Johannes Muhle-Karbe, ETH Zürich

Wien, im Jänner 2011

Abstract

This thesis analyzes pricing and hedging methods for various volatility derivatives.

In the first section, we essentially only assume that the stock price process is continuous and we show that the information in European option prices contains the price of a generalized variance swap. We also show how to hedge this swap by means of robust trading strategies. Finally, we analyze the impact of jumps in the underlying stock price process on our formulas.

For our analysis of the volatility swap, we make additional assumptions on the stock price process and we assume that the volatility process and the stock price process are independent of each other. Under these conditions, we can find the price of the volatility swap in terms of Europeans. We then make a numerical analysis of the impact of jumps.

In the final chapter, we find subreplication strategies for the variance call assuming only continuity for the stock price. Ultimately, we find a Black Scholes style formula under additional assumptions and test it numerically with a model with jumps.

Zusammenfassung

Diese Diplomarbeit beschäftigt sich mit Methoden zur Preisbestimmung und zur Replikation von Volatilitätsderivaten.

Für den Variance Call lässt sich eine Replikationsstrategie finden, welche in sämtlichen stetigen Modellen gültig ist. Wir zeigen eine Formel mit welcher wir den Preis des Variance Calls mit Europäischen Call und Put Optionspreisen berechnen könnnen. Zustäzlich analysieren wir den Effekt von Sprüngen des Aktienprozesses für unsere Formeln.

Im zweiten Abschnitt machen wir zusätzliche Annahmen an den Aktienpreisprozess und verlangen, dass der Aktienpreisprozess und der Volatilitätsprozess unabhängig sind. Unter diesen Vorraussetzungen finden wir eine Formel für den Volatility Swap. Anschließend untersuchen wir numerisch wie stark der Fehler unserer Formel ist, wenn der Aktienprozess Sprünge hat.

Im letzten Abschnitt untersuchen wir den Variance Call und finden Strategien dessen Wert zum Verfallstag höchstens gleich dem des Variance Calls ist. Danach untersuchen wir numerisch eine besonders einfache Formel zur Berechnung des Preises des Variance Calls, welche aber nur unter zusätzlichen Annahmen gültig ist.

Acknowledgments

I would like to thank my thesis advisors Ass. Prof. Johannes Muhle-Karbe, who introduced me to the topic, and Prof. Walter Schachermayer, who raised my interest in mathematical finance with his excellent lecture courses. I am grateful for their support, patience and guidance.
Furthermore, I wish to thank my family and friends for their support during my studies.

Contents

1. Variance swaps

The first and simplest volatility derivative that we are going to analyze in this thesis is the variance swap. Let us define the process of the log returns as

$$X_T := \log(S_T/S_0)$$

where S_T denotes the stock prices process. A variance swap pays at some fixed terminal time T the amount

$$[X]_T - K \tag{1}$$

where $[.]_t$ denotes the quadratic variation and K is some fixed agreed-upon amount. Throughout this thesis, we assume, for simplicity that $K = 0$ and that the interest rates are zero. Working in a very general framework, we essentially only assume continuity of the stock price process.

In this first chapter, we mainly follow Carr and Lee [7].

1.1. Assumptions.

Let $T > 0$ be a finite terminal time.

Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space. We assume that our filtration $\mathcal{F}_t$ is right continuous, i.e. :

$$\bigcap_{u>t} \mathcal{F}_u = \mathcal{F}_t$$

and that $\mathcal{F}_0$ is the trivial σ-field augmented by the $\mathbb{P}$-null sets.

We assume that our stock price process S_t is a positive continuous semi-martingale adapted to $\mathcal{F}_t$ with the initial condition $S_0 \in \mathbb{R}_+$.

Furthermore, we assume that there exists an equivalent risk neutral measure $\mathbb{Q}$ such that S_t is a martingale under $\mathbb{Q}$.

Remark In this thesis we analyze volatility contracts which depend on the realized variance or realized volatility, in contrast to volatility derivatives which depend on e.g. implied volatility.

In practice a variance swap pays for some partition $\{0 = t_0 \leq t_1 ... \leq t_n = T\}$:

$$\sum_{i=0}^{n-1} \log(S_{t_{i+1}}/S_{t_i})^2. \tag{2}$$

The term converges in probability to the quadratic variation of X_T as the mesh of the partition $\{0 = t_0 \leq t_1 ... \leq t_n = T\}$ tends to zero. Hence we already have an idealization in the definition of our variance swap, when we use the quadratic variation.

Definition 1.1. *Let ϕ:$(0, \infty) \to \mathbb{R}$ be a difference of convex functions, we then define the process X_t^ϕ as*

$$X_t^\phi := \phi(S_t) \tag{3}$$

and for the special case $\phi = \log(S_t/S_0)$

$$X_t := \log(S_t/S_0). \tag{4}$$

Furthermore we assume that:

$$\mathbb{E}[[X]_T] < \infty. \tag{5}$$

Remark Note that throughout this thesis we will only work under the risk neutral measure. Every expectation is therefore, to be understood as one under the risk neutral measure.

Of course, our assumptions include all the typical stock price process, such as local-volatility or stochastic volatility models, where the price process has no jumps. The stock price process and the volatility process can also be correlated.

1.2. **Replicating variance swaps.** To begin with, we need some results on convex functions. Recall that a function is convex on an open intervall $I \subset \mathbb{R}$ if :

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \qquad \forall x, y \in I, \in [0,1].$$

The next lemma provides the properties of convex functions that we need.

Lemma 1.2. *Let f be a a convex function defined on an open interval $I \subset \mathbb{R}$.*

Then, at each point $x \in I$ the left-hand derivative $f'_-(x)$ and the right-hand derivative $f'_+(x)$ exist.

The functions $f'_-(x)$ and $f'_+(x)$ are increasing,left- resp. right-continuous and the set $\{x : f'_-(x) \neq f'_+(x)\}$ is at most countable and where $f'_- = f'_+$, the function f'_- is continuous.

Furthermore, the second derivative f'' exists in the sense of distributions and it is a positive measure.

Proof. See Revuz and Yor [17] (Appendix 3.1 and 3.2). ☐

Note that the above lemma implies $\int_0^t f'_+(x)\,dx = \int_0^t f'_-(x)\,dx$.

The next theorem will generalize Itô's formula. Instead of demanding that $f \in C^2$, we only assume that f is a difference of convex functions. For a function $f \in C^2$ one can set $f_+ = f\mathbf{1}_{f'' \geq 0}$ and $f_- = -f\mathbf{1}_{f''<0}$ and one can write $f = f_+ - f_-$ as a difference of convex functions.

Theorem 1.3. *Let M be a continuous semi-martingale. If f is a difference of convex functions we have:*

$$f(M_t) = f(M_0) + \int_0^t f'_-(M_s)\,dM_s + \frac{1}{2}\int_{-\infty}^{\infty} L_t^a f''(da) \tag{6}$$

where L_t^a denotes the local time of M, which is an increasing continuous process.

Proof. See Revuz and Yor [17] (Theorem VI.1.5). □

Additionally, we will need the following result.

Proposition 1.4. *Under the assumptions of the above theorem, we have a.s.*

$$\int_0^t g(M_s)\, d[M]_s = \int_{-\infty}^{\infty} g(a) L_t^a\, da \qquad \forall t \in \mathbb{R}_+ \tag{7}$$

where g is a positive Borel function.

Proof. See Revuz and Yor [17] (Corollary VI.1.6). □

To better understand the above formula let us set $g = \mathbf{1}_A$ and $M_t = B_t$, where B_t is a Brownian Motion. Then we have $\int_{-\infty}^{\infty} \mathbf{1}_A L_t^a\, da = \int_0^t \mathbf{1}_A\, dt$. We see that the left-hand side represents the amount of time of the Brownian motion spent in A. Therefore, the formula is referred to as the "Occupation times formula".

We will first prove a more general result, from which we get the result for the interesting variance swap as a special case.

The next definition generalizes the variance swap.

Definition 1.5. *We say ω is a weight function if ω is a Borel measurable function and $\omega : (0,\infty) \to [0,\infty)$.*
Let $\phi : (0,\infty) \to \mathbb{R}$ be a difference of convex functions s.t. $\mathbb{E}[[X^\phi]_T] < \infty$ and let τ be a stopping time. Then define the forward starting weighted variance swap of $\phi(S)$ as

$$[X^\phi]^w_{\tau,T} := \int_{\tau\wedge T}^{T} \omega(S_u) d[X^\phi]_u. \tag{8}$$

By (5) the log-case is included in the above definition.

Let us apply Theorem 1.3 to our process X_t^ϕ. We then have

$$X_t^\phi = X_0^\phi + \int_0^t \phi_{y-}(S_t)\, dS_t + \frac{1}{2}\int_{-\infty}^{\infty} L_t^a \phi_{yy}(da). \tag{9}$$

where ϕ_{y-} denotes the left-hand derivative with respect to y.

Suppose the measure ϕ_{yy} has a density, i.e. there exists a Borel function $f \geq 0$ s.t. $\int_A f\, dy = \phi_{yy}(A)$ for all Borel sets A, then we have

$$\frac{1}{2}\int_{-\infty}^{\infty} L_t^a \phi_{yy}(da) = \frac{1}{2}\int_{-\infty}^{\infty} L_t^a f(a)\, da = \frac{1}{2}\int_0^t f(S_u)\, d[S]_u \qquad a.s. \tag{10}$$

by Proposition 1.4, where L_t^a is the local time of S. Thus, we see that in this case we have the Itô formula in the usual form. Since the term on the right hand side is of finite variation, we get

$$d[X^\phi]_t = \phi_{y-}^2(S_t)d[S]_t. \tag{11}$$

The next theorem shows how to replicate our forward starting weighted variance and it is the main result of this section.

Theorem 1.6. *Let ω be a weight function and let τ be a stopping time. Let ϕ be as in Definition 1.5 and let $\chi : (0,\infty) \to [0,\infty)$ be a difference of convex functions and assume that it satisfies $\forall y \in (0,\infty)$*

$$\chi_{yy}(y) \leq (\geq) 2\phi_{y-}^2(y)\omega(y) \tag{12}$$

and

$$\mathbb{E}[\chi(S_T)] < \infty \text{ as well as } \mathbb{E}[\chi(S_{T\wedge\tau})] < \infty. \tag{13}$$

Furthermore, we assume that the signed measure of the second derivative of χ has a density denoted by χ_{yy}.

Then the following strategy subreplicates (superreplicates) the forward starting weighted variance of $\phi(S)$.

At each time $t \in (0, T \wedge \tau]$ hold:

1 claim on $\chi(S_T)$

1 claim on $-\chi(S_{\tau\wedge T})$

and at each time $t \in (\tau \wedge T, T]$ hold:

1 claim on $\chi(S_T)$

-$\chi_{y-}(S_t)$ shares

$\chi(S_{\tau\wedge T}) - \int_{\tau\wedge T}^t \chi_{y-}(S_u)\, dS_u + S_t\chi_{y-}(S_t)$ bonds

The value V_0 of our subreplicating (superreplicating)portfolio at time 0 is given by

$$V_0 = \mathbb{E}[\chi(S_T)] - \mathbb{E}[\chi(S_{\tau\wedge T})]. \tag{14}$$

If we have equality in (12) *the strategy replicates the forward starting weighted variance exactly.*

Proof. We prove the case of the subreplicating portfolio and look at the value of the portfolio at time T. Let us apply Theorem 1.3 and (10):

$$\chi(S_T) = \chi(S_0) + \int_0^T \chi_{y-}(S_u)\, dS_u + \int_0^T \frac{1}{2}\chi_{yy}(S_u)\, d[S]_u$$

for the stopped process we get

$$\chi(S_{\tau\wedge T}) = \chi(S_0) + \int_0^{\tau\wedge T} \chi_{y-}(S_u)\, dS_u + \int_0^{\tau\wedge T} \frac{1}{2}\chi_{yy}(S_u)\, d[S]_u$$

We combine this with (12) to get:

$$\begin{aligned}\chi(S_T) &= \chi(S_{\tau\wedge T}) + \int_{\tau\wedge T}^{T} \chi_{y-}(S_u)\, dS_u + \int_{\tau\wedge T}^{T} \frac{1}{2}\chi_{yy}(S_u)\, d[S]_u \\ &\le \chi(S_{\tau\wedge T}) + \int_{\tau\wedge T}^{T} \chi_{y-}(S_u)\, dS_u + \int_{\tau\wedge T}^{T} \phi_{y-}^2(S_u)\omega(S_u)\, d[S]_u \\ &= \chi(S_{\tau\wedge T}) + \int_{\tau\wedge T}^{T} \chi_{y-}(S_u)\, dS_u + \int_{\tau\wedge T}^{T} \omega(S_u)\, d[X^{\phi}]_u\end{aligned}$$

where we used $\phi_{y-}^2 d[S]_u = d[X^{\phi}]_u$ in the last equation. Therefore, for the payoff of our strategy at time t we get the following inequality by the definition of the forward starting weighted variance:

$$\chi(S_T) - \chi(S_{\tau\wedge T}) - \int_{\tau\wedge T}^{T} \chi_{y-}(S_u)\, dS_u \le [X^{\phi}]^{w}_{\tau,T}.$$

We see that we get perfect replication when we have an equality in (12). That the strategy has the claimed time zero value and that it is self-financing is obvious. □

We see that we start the trading in the stock as soon as $[X^{\phi}]^{\omega}_{\tau,T}$ starts running. Suppose that we have for the stopping time $\tau \ge T$ then we get $[X^{\phi}]^{\omega}_{\tau,T} = 0$. In our subreplicating portfolio we have one claim on $\chi(S_T)$, one on $-\chi(S_{\tau\wedge T})$ and no share position. This portfolio pays zero at time T and hence the value is equal to the forward starting weighted variance.

Remark For this result to be practical we need to be able to buy and sell claims on $\chi(S_T)$ and $-\chi(S_{\tau\wedge T})$ in the market. This assumption might not be satisfied.

Obviously, the above theorem includes replication of the variance swap as a special case.

Corollary 1.7. *We have the following self-financing strategy for the variance swap, at each time $t \le T$ hold:*

-2 log contracts which pay $\log(S_T/S_0)$

$2/S_t$ *shares*

$\int_0^t 2/S_u\, dS_u - 2$ *bonds*
The strategy has time zero value:

$$\mathbb{E}[[X]_T] = \mathbb{E}[-2\log(S_T/S_0)]. \tag{15}$$

At general times t we have:

$$\mathbb{E}[[X]_T|\mathcal{F}_t] = [X_t] + \mathbb{E}[-2\log(S_T/S_0) + 2(S_T/S_t) - 2|\mathcal{F}_t]. \tag{16}$$

Proof. We choose in Theorem 1.6 $\tau = 0$, $\chi(y) = -2\log(y/S_0)$ and $\omega = 1$. The value for the portfolio at general times is trivial. □

We see that in order to replicate a variance swap, we need to hold the static amount of 2 dollars in our stock by continuous trading.

Example Suppose we want to exactly replicate the "gamma swap" with the payoff

$$\int_{\tau \wedge T}^{T} aS_u \, d[X]_u$$

where $a \in \mathbb{R}$. That means we choose in Theorem 1.6 $\omega = ay$. Let $\phi(y) = \log(y)$ and we have by Theorem 1.6 the ODE for χ

$$\chi_{yy}(y) = 2a/y$$

which we can easily solve and get the solution

$$\chi(y) = ay\log(y) + by + c$$

for some constants b, c. Thus, we have found the possible claims χ of Theorem 1.6 we can hold to replicate the "gamma swap".

1.3. **Time change.** We derive the price for the variance swap a second time with a time change, which gives us further insight.

Definition 1.8. *A time change C is a family $\{C_t : t \geq 0\}$ of stopping times with the property that $t \mapsto C_t$ is right continuous and increasing.*

A particularly easy example of a time change is given by $C_t = t \wedge \tau$.

We will need the following classic theorem.

Theorem 1.9. *Let M be continuous local martingale adapted to $\mathcal{F}_t$. Let $M_0 = 0$ and let $[M, M]_\infty = \infty$. We set:*

$$k_t = \inf\{s : [M]_s > t\}$$

Then we have:
$B_t = M_{k_t}$ is a $\mathcal{F}_{k_t}$ Brownian Motion, k_t is a time change and that

$$M_t = B_{[M]_t}. \tag{17}$$

Proof. See Revuz and Yor [4] (Theorem V 1.6). □

Definition 1.10. *Define for a semi-martingale M*

$$[M]_\infty := \lim_{t \to \infty} [M]_t. \tag{18}$$

The next theorem will give us additional information about the development of the value of the variance swap. We generalize the proof from Carr, Lee and Wu [8](Proposition 1.4), where the case $\phi(S_t) = \log(S_t)$ is treated.

Theorem 1.11. *Let S satisfy our usual assumptions and let ϕ be as in Definition 1.5. Let χ be a difference of convex functions with density function χ_{yy} for the measure of the second derivative. Furthermore assume that*

$$\chi_{yy}(y) = \phi^2_{y-}(y) \tag{19}$$

and

$$\mathbb{E}[\chi(S_T)] < \infty \tag{20}$$

as well as

$$[\chi(S)]_\infty = \infty. \tag{21}$$

Then there exists a filtration $\mathcal{G}_u$ and a continuous time change C such that $\mathbb{E}[C_T] < \infty$ such that

$$\phi(S_t) = W_{C_t} - C_t/2$$

where W is a $\mathcal{G}$ Brownian motion.

Proof. By Theorem 1.3 we have:

$$\chi(S_t) = \int_0^t \chi_{y-}(S_u)\, dS_u + \frac{1}{2}\int_0^t \chi_{yy}(S_u)\, d[S]_u = \int_0^t \chi_{y-}(S_u)\, dS_u + \frac{1}{2}\int_0^t d[X^\phi]_u$$

where we used $d[X^\phi]_u = \phi^2_{y-}(S_u)d[S]_u = \chi_{yy}(S_u)d[S]_u$.
We have $\int_0^t d[X^\phi]_u = [X^\phi]_t$, by the chain of rule of the Riemann Stieltjes integral, we then set

$$M_t := \int_0^t \chi_{y-}(S_u)\, dS_u = \chi(S_t) + \frac{1}{2}[X^\phi]_t.$$

M_t is a continuous martingale since the Itô integral with respect to a continuous martingale is again a continuous martingale.
Let us set $k_u := \inf\{t : [M]_t \geq u\}$, $\mathcal{G}_u := \mathcal{F}_{k_u}$ and $C_t := [M]_t = [X^\phi]_t$, where we use that $[X^\phi]_t$ is of finite variation. We use Theorem 1.9, which yields that $W_u := M_{k_u}$ is a $\mathcal{G}$ Brownian motion and $W_{C_t} = M_t$. Hence we get

$$\chi(S_t) = W_{C_t} - C_t/2.$$

□

If we now take the expectation of $\chi(S_t) = W_{[X^\phi]_t} - \frac{[X^\phi]_t}{2}$ we arrive back at our result:

$$\mathbb{E}[[X^\phi]_T] = \mathbb{E}[-2\chi(S_t)]$$

where an obvious choice for $\chi(y)$ is $\log(y/S_0)$.

Note that for this more elegant derivation we needed our assumption that we have a risk neutral measure. To prove our hedging strategy we could even drop this assumption.

Remark The technique of time change that we used in this section becomes an essential tool if one wants to value the variance swap when the underlying process is a Lévy process with jumps. In [8] Carr, Lee and Wu consider a process for the log returns of the form X_{C_t}, where X is a non-constant Lévy process and C_t is a time change. They show that for a time change C_T such that $\mathbb{E}[C_T] < \infty$ we get:

$$\mathbb{E}[[X]_T] = Q_S \mathbb{E}[-\log(S_T/S_0)] \tag{22}$$

where the multiplier Q_S is some constant which depends on the underlying Lévy process X_t but not on the time change C_t. For the special case where S_t is a continuous martingale, we have shown that $Q_S = 2$. If we assume that our stock price process is of the form $S_t = e^{X_t}$ where X_t is a $NIG(\alpha, \beta, \delta)$ process it is shown in Carr, Lee and Wu [8] that the multiplier Q_S is of the form

$$Q_S = \frac{\alpha^2/(\alpha^2 - \beta^2)}{\alpha^2 - \beta^2 - \beta - \sqrt{(\alpha^2 - \beta^2)(\alpha^2 - (\beta+1)^2)}}. \tag{23}$$

When we increase β, we increase the size of the up-jumps and we see that with higher up-jumps our multiplier decreases. On the other hand, increasing the sizes of down jumps (decreasing β) results in a higher multiplier.
If the stock price is modeled via a VG process with parameters C, G, M Carr-Lee show that

$$Q_S = \frac{1/G^2 + 1/M^2}{1/G - \log(1 + 1/G) - 1/M - \log(1 - 1/M)}. \tag{24}$$

Here we see that the parameter C is irrelevant to the multiplier.
In the VG-model we have larger up-jumps if $G > M$. In this case, we see that we again get a smaller multiplier. As before we have larger down-jumps with $G < M$, which leads to a larger multiplier.

We will consider the NIG and the VG processes again, when we analyze the impact of jumps for volatility swaps where we provide some basic properties of these processes.

1.4. **Pricing variance swaps in terms of Europeans.** We have seen that the problem of valuing the variance swap under our assumption is equivalent to the task of valuing the log contract. In the following, we will prove that we can price the log contract model independent from Europeans, which then also yields the value of the variance swap. We will need the following theorem.

Theorem 1.12. *Let $G : \mathbb{R}_+ \to \mathbb{R}$ be a difference of convex functions where the second derivative in the distributional sense has a density G''.*
Then for any $F \in \mathbb{R}$ and for all $X \in \mathbb{R}_+$:

$$G(X) = G(F) + G'_-(F)(X - F) + \int_0^F G''(K)(K - X)^+ \, dK$$

$$+ \int_F^\infty G''(K)(X-K)^+\,dK.$$

Proof. For a sketch of a proof see Carr-Madan [5] (Appendix). $\square$

We recognize the first two terms as the Taylor approximation at F and the integral terms continuously bend the right hand side to the nonlinear function.

We can now use this Theorem to price our (sub)replicating strategy of Theorem 1.6 in terms of European call and put options.
For the important special case of the ordinary variance swap, we have the following corollary.

Corollary 1.13. *Let $C(K)$, $P(K)$ denote the risk neutral prices of call/put options with strike K and time zero stock price S_0 and time to expiry T.
Let χ be a difference of convex functions where the second derivative in the distributional sense has a density χ_{yy}.
We then have*

$$\begin{aligned}\mathbb{E}[\chi(S_T)] =& \chi(F) + \chi'_-(F)(C(F)-P(F)) \\ &+ \int_0^F \chi_{yy}(K)P(K)\,dK + \int_F^\infty \chi_{yy}(K)C(K)\,dK \end{aligned} \tag{25}$$

and for our variance swap

$$\mathbb{E}[[X]_T] = 2\int_0^{S_0} \frac{1}{K^2}P(K)\,dK + 2\int_{S_0}^\infty \frac{1}{K^2}C(K)\,dK. \tag{26}$$

Proof. For the first equation simply take the expectation in Theorem 1.12.
By Theorem 1.12 we have

$$\log(S_T/S_0) = -\int_0^{S_0} \frac{1}{K^2}(K-X)^+\,dK - \int_{S_0}^\infty \frac{1}{K^2}(X-K)^+\,dK.$$

The term with the first derivative cancels because we choose F of Theorem 1.12 to be S_0 and we use the martingale property of S_t. Next we use (15) and take the expectation to get our desired result. $\square$

Remark Theoretically, to use the above result to price a variance swap we need the prices of put and call options for all strikes. Of course, in practice we do not have this information. However, in liquid markets we usually have enough prices to make useful approximations of our integral. Moreover, we now also know how to hold the log-contract from Corollary 1.7 and the claim on $\chi(S_T)$ from Theorem 1.6 , by holding the above option position. Note that the position is static because the terms in the integral in (25) and (26) only depend on the strike K and not on S_t or t. This is essential since hedging strategies where the continuous trading of options is necessary are not practical due to transaction costs.

Note that for the replication of the forward starting variance swap we also need to hold a claim on $-\chi(S_{\tau\wedge T})$. We can not replicate this claim with Europeans. Thus, for the forward starting variance swap ($\tau \neq 0$) we still have to assume that $-\chi(S_{\tau\wedge T})$ trades.

Example As a sanity check let in $\chi(S_T) = (S_T - Q)^+$ the above corollary. Then in the distributional sense we have $\chi_(y)' = \mathbf{1}_{(y>Q)}$ and $\chi(y)'' = \delta(y-Q)$ and we get

$$\mathbb{E}[(S_T - Q)^+] = (S_0 - Q)^+ + \int_0^{S_0} P(K)\delta(K-Q)\,dK + \int_{S_0}^{\infty} C(K)\delta(K-Q)\,dK.$$

If $S_0 \geq Q$ the right-hand side obviously equals $C(Q)$.
If $S_0 < Q$ we have for the right-hand side $(S_0 - Q) + P(Q)$ and we use the put-call parity $C(K) - P(K) = S_0 - K$ to obtain $C(Q)$.

That means that if we want to replicate a European call option with European options we need to hold this particular European call option.

Remark We are aware of the fact that we are not being fully rigorous in this example since the Dirac delta measure has no density function.

Example For the "arithmetic variance swap" we have in Theorem 1.6 $\phi(y) = y$, $\omega = 1$ and $\tau = 0$. Then we choose some $a \in \mathbb{R}$ and we set $\chi(y) = (y-a)^2$. Then we have $\chi_{yy} = 2\phi_y^2(y)\omega(y)$ and therefore Theorem 1.6 guarantees replication. We have by (25):

$$\mathbb{E}[\chi(S_T)] = \int_0^a 2P(K)\,dK + \int_a^{\infty} 2C(K)\,dK \tag{27}$$

and this is the integral we have to approximate to price this derivative.

Example Let B be a Borel set. If we choose in Theorem 1.6 $\phi(y) = \log(y/S_0), \omega(y) = \mathbf{1}_{y\in B}$ and $\tau = 0$, we have the "corridor variance swap". The corresponding χ of Theorem 1.6 is given by $\chi(y) = -2\log(y/S_0)\mathbf{1}_{y\in B}$ and we can approximate it's price by:

$$\mathbb{E}[\chi(S_T)] = \int_{\{K\geq S_0 : K\in B\}} \frac{2}{K^2}C(K)\,dK + \int_{\{K<S_0 : K\in B\}} \frac{2}{K^2}P(K)\,dK \tag{28}$$

where we used (25).

Example Consider the Heston model (Heston [14]) given by:

$$dS_t = \sqrt{v_t}S_t dW_t^1$$

$$dv_t = \kappa(\theta - v_t)dt + \xi v_t W_t^2$$

where $\kappa \in \mathbb{R}, \theta, \xi \in \mathbb{R}_+$ and dW^1 and dW^2 are Brownian motions which can be correlated:

$$[dW_t^1, dW_t^2] = \rho dt$$

where $\rho \in [-1, 1]$.
In this model we can find an explicit formula for our variance swap (See Gatheral [13] page 138)

$$\mathbb{E}[X_T] = \frac{1 - e^{-\kappa T}}{\kappa}(v_0 - \theta) + \theta.$$

Note that the expected variance does not depend on the volatility of the volatility. This implies that while a single European option depends on the volatility of the volatility (ξ turns up in the European option price formula of the Heston model), our strip of European options does not.

1.5. **The impact of jumps.** In this subsection, we follow Gatheral [13] (chapter 10).

At first we will show how to value the variance swap in a pure jump model. We assume that our log-return process $P_t := \log(S_t/S_0)$ is a compound Poisson process.
Let y_i be identically distributed pairwise independent processes with density μ and let N_t be a Poisson process with intensity λ. We also assume that y_j and N_t are independent and that $\mathbb{E}[y_i] = v < \infty$. We set P_t:

$$P_t = \sum_{j=1}^{N_t} y_j. \tag{29}$$

We then have for the quadratic variation:

$$[P]_t = \sum_{j=1}^{N_t} |y_j|^2. \tag{30}$$

We have

$$\mathbb{E}[[P]_T] = \mathbb{E}[N_T]\mathbb{E}[|yi|^2] = \lambda T \int_0^\infty t^2 \mu(t)\, dt$$

where we used the independence of N_t and y_i for the first equality. The second equality follows from the well-known results of the Poisson process and the expectation.
We compute the variance of our compound Poisson process and get:

$$\begin{aligned}
Var(P_t) &= \mathbb{E}[Var(P_t|N_t)] + Var(\mathbb{E}[P_t|N_t]) \\
&= \mathbb{E}[N_t Var(y_1)] + Var(N_t \mathbb{E}[y_1]) \\
&= Var(y_1)\mathbb{E}[N_t] + \mathbb{E}[y_1]^2 Var(N_t) \\
&= Var(y_1)\lambda t + \mathbb{E}[y_1]^2 \lambda t \\
&= \lambda t (Var(y_1) + \mathbb{E}[y_1]^2) \\
&= \lambda t \mathbb{E}[y_1^2] \\
&= \lambda T \int_0^\infty t^2 \mu(t)\, dt
\end{aligned}$$

where the first equality follows from the law of total variance. The second inequality follows by independence of the y_i. The other inequalities use elementary properties of the variance and the fact that the Poisson process at time t is $P(\lambda t)$ (where P denotes the Poisson distribution) distributed which implies $Var(N_t) = \lambda$.

We have shown that if our log return process is a compound Poisson process, then the value of our variance swap equals the variance of our process. We can use Theorem 1.12 to price the variance swap in terms of Europeans.

We then have for our price, where we again let $C(K), P(K)$ denote our risk neutral call/put prices:

$$
\begin{aligned}
\mathbb{E}[[P]_T] &= Var(P_T) \\
&= \mathbb{E}[P_T^2] - \mathbb{E}[P_T]^2 \\
&= \int_0^{S_0} \frac{2}{K^2} - \frac{2\log(K/S_0)}{K^2} P(K)\,dK + \int_{S_0}^{\infty} \frac{2}{K^2} - \frac{2\log(K/S_0)}{K^2} C(K)\,dK \\
&- [-\int_0^{S_0} \frac{1}{K^2} P(K)\,dK - \int_{S_0}^{\infty} \frac{1}{K^2} C(K)\,dK]^2
\end{aligned}
$$

where we use (25) for $\mathbb{E}[\log(S_T/S_0)^2]$ and $\mathbb{E}[\log(S_T/S_0)]^2$.

This formula shows that assuming that the log return process is a compound Poisson process, we can, as in the continuous case, get the price of the variance swap out of European call and put options. However, comparing the formulas shows that the weights of the strips of the Europeans do not coincide. Hence, in this framework we get a different formula for the price of the variance swap compared to the continuous case.

Since in reality we do not know whether the stock price process is continuous or not, we are interested in the error we that we get when pricing the variance swap with the log strategy (which holds true in continuous models) in the presence of jumps. Therefore we want to calculate the difference between the true value of the variance swap and the value of $-2\mathbb{E}[\log(\frac{S_T}{S_0})]$ in jump models.

Suppose our log returns process is a sum of continuous semi-martingale and a compound Poisson process, where the jump and the continuous part are independent of each other. In this case we can compute this difference. In Gatheral [13](chapter 10) this is done by using the Lévy-Khintchine representation, it is shown that:

$$
\mathbb{E}[[X]_T] - (-2\mathbb{E}[X_T]) = 2\lambda T \int_0^T (1 + t + t^2/2 - e^t)\mu(t)\,dt. \tag{31}
$$

We can calculate the impact of jumps exactly in the case when the jumps sizes are log-normally distributed, i.e. we assume for our jump sizes at time

t the following distribution:

$$(e^{\alpha+\delta\epsilon}-1)S_{t-} \tag{32}$$

where ϵ is $N(0,1)$ distributed.
Then, using (31), we calculate:

$$\mathbb{E}[[X]_T]-(-2\mathbb{E}[X_T]) = \lambda T(\alpha^2+\delta^2)+2\lambda T(1+\alpha+e^{\alpha+\delta^2/2}). \tag{33}$$

From this formula, we obtain that when we increase the size of down-jumps the log contract increasingly underestimates the price. Whereas when increase the size of up-jumps, it overestimates the value. We have seen the same pattern for the multipliers of the jump-models presented in the remark before. Carr-Lee show that in a sense this holds in general (see [8] Proposition 4.3).

We calculate the error with some calibrated parameters from the literature (setting $T = 2$):

Reference	λ	α	δ	error
Andersen and Andreasen[1]	0.59	-0.05	0.07	0.37%
Bakshi, Cao and Chen[2]	0.5	-0.15	0	1.07%
Duffie, Pan, and Singleton[11]	0.11	-0.12	0.15	0.75%
Jim Gatheral[13]	0.13	-0.12	0.10	0.51%

From this table and from the formula above, we find that for reasonable parameters the error of using the log strategy when the stock price really has log-normally distributed jumps is negligible. The errors are given in terms of the true value.

2. Volatility swaps

The next simplest volatility derivative that we are going to consider is the volatility swap. A volatility swap pays at time T

$$\sqrt{[X]_T} - K$$

where K is again a fixed agreed-upon amount which we take to be zero for simplicity. We would like to obtain similar results for this payoff as we got for the variance swap. Thus, the goal of this section is to relate $\mathbb{E}[\sqrt{[X]_T}]$ to European option prices. In this section we mainly follow Carr and Lee [6].

2.1. **Assumptions.** In order to price the volatility swap we have to make more assumptions on our underlying stock price process than in the variance swap case.
We assume that $S_0 > 0$ and

$$dS_t = \sigma_t S_t dW_t \tag{34}$$

where W_t is a Brownian Motion adapted to $\mathcal{F}_t^W \subset \mathcal{F}_t$ and σ_t is adapted to $\mathcal{F}_t^\sigma \subset \mathcal{F}_t$.

Furthermore, we assume that the volatility process σ_t is cádlág (i.e. right continuous where the left-hand limits exist) and independent of W_t, i.e. for every t the σ-algebras $\mathcal{F}_t^W$ and $\mathcal{F}_t^\sigma$ are independent

and

$$\mathbb{E}[[X]_T] < \infty. \tag{35}$$

Under these assumptions our stock price process exists and is of the form:

$$S_t = S_0 e^{\int_0^t \sigma_u \, dW_u - \frac{1}{2}\int_0^t \sigma_u^2 \, du} \tag{36}$$

and we have

$$[X]_t = \int_0^t \sigma_u^2 \, du. \tag{37}$$

Remark Note that we still do not make any assumption on the particular process that the volatility follows, which makes our theory valuable. Our assumptions include stochastic volatility models as well as local volatility models, in contrast to the stock price process the volatility is also allowed to jump. Of course the independence assumption is completely unrealistic in equity markets. When the stock price goes down, investors tend to get nervous and the volatility goes up, this means that the process are typically chosen to be negatively correlated. In this thesis we will not try to weaken this assumption.

Remark Under these assumptions we can find an elegant formula for the value of the variance swap of the previous section. Let $k := \log(K/S_0)$

denote the log strike and let

$$c(k) := \frac{C(S_0 e^k, T, S_0)}{S_0 e^k} \text{ and } p(k) := \frac{P(S_0 e^k, T, S_0)}{S_0 e^k}. \tag{38}$$

Under our independence assumption,we have, by Carr and Lee [10](Corollary 2.5) the following put-call symmetry(not pararity!):

$$\mathbb{E}[(S_T - K)^+] = \frac{K}{S_0}\mathbb{E}[(\frac{S_0^2}{K} - S_T)^+]$$

which then gives

$$e^{-k}c(-k) = \frac{S_0}{K}\frac{C(\frac{S_0^2}{K}, T, S_0)}{S_0 e^k} = p(k).$$

We use this equality to obtain

$$\begin{aligned}
\mathbb{E}[[X]_T] &= 2\int_0^{S_0} \frac{1}{K^2}P(K)\,dK + 2\int_{S_0}^{\infty} \frac{1}{K^2}C(K)\,dK \\
&= 2\int_{-\infty}^{0} p(k)\,dk + 2\int_0^{\infty} c(k)\,dk \\
&= 2\int_0^{\infty} c(k)e^k\,dk + 2\int_0^{\infty} c(k)\,dk \\
&= 4\int_0^{\infty} c(k)e^{k/2}\cosh(k/2)\,dk
\end{aligned}$$

where we use $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$.We have found a formula for the variance swap which depends only on European call prices and we need to numerically evaluate only one integral.

2.2. **Boundary Values and Approximations.** In this section we will find boundary values and approximations for the price of our volatility swap.

Definition 2.1.

Let $BS_{atmc}(S_0, \sigma)$ denote the at the money price of a European call option with initial stock price value S_0 in the Black-Scholes model with volatility σ.

Define IV_0 as the unique solution to

$$BS_{atmc}(S_0, IV_0) = \mathbb{E}[(S_T - S_0)^+].$$

Remark Of course, we know that $BS_{atmc}(S_0, \sigma) = S_0(N(\sigma/2) - N(-\sigma/2))$.

First we prove a useful boundary value for the Black-Scholes implied volatility.

Proposition 2.2. *Under the assumptions of this section, without needing the independence assumption, we have for IV_0:*

$$\frac{\sqrt{2\pi}}{S_0}\mathbb{E}[(S_T - S_0)^+] \leq IV_0. \tag{39}$$

Proof. By a simple calculation of $(S_0(N(\sigma/2) - N(-\sigma/2)))''$, we see that $BS_{atmc}(S_0, .)$ is concave. Hence, it must lie below it is tangent at zero which is given by $\sigma \to \frac{S_0\sigma}{\sqrt{2\pi}}$. Thus, we have:

$$\frac{\sqrt{2\pi}}{S_0}\mathbb{E}_0(S_T - S_0)^+ = \frac{\sqrt{2\pi}}{S_0} BS_{atmc}(S_0, IV_0) \leq \frac{\sqrt{2\pi}}{S_0}\frac{S_0 IV_0}{\sqrt{2\pi}}.$$

□

Next we find an upper bound for the value of our volatility swap :

Proposition 2.3. *Under the assumptions of this section :*

$$IV_0 \leq \mathbb{E}[\sqrt{[X]_T}] \leq \sqrt{\mathbb{E}[[X]_T]}. \tag{40}$$

Proof. For a proof of the first inequality see Carr and Lee [6](Prop. 6.1). For the second inequality, we use the conditional Jensen's Inequality, with the convex function $y \to -\sqrt{y}$. □

These approximations for the value of the volatility swap can be very good, especially under our standing zero-correlation assumption, see Carr and Lee [6] (Remark 6.3-6.5) for further details.

2.3. **Pricing volatility swaps.** In this section we will find exact formulas, for the price of our volatility swap. We will need the following two propositions.
We need the Gamma function which is defined for $z > 0$ by

$$\Gamma(z) := \int_0^\infty u^{z-1}e^{-u}\, du. \tag{41}$$

It is well known that the integral converges.
We will need the results of the next proposition later on.

Proposition 2.4. *Let X be a random variable with $X \geq 0$.*
Let $\varphi_X(t) := \mathbb{E}[e^{-tX}]$ denote the Laplace transform of X and assume that $\mathbb{E}[X] < \infty$. We have for $0 < r < 1$:

$$\mathbb{E}[X^r] = \frac{r}{\Gamma(1-r)}\int_0^\infty \frac{1-\varphi_X(t)}{t^{r+1}}\, dt \tag{42}$$

and for $p > 0$

$$\mathbb{E}[X^{-p}] = \frac{1}{\Gamma(p)p}\int_0^\infty \varphi_X(t^{1/p})\, dt. \tag{43}$$

Proof. We have by the definition of the Γ-function for $0 < r < 1$

$$\Gamma(1-r) = \int_0^\infty \frac{e^{-u}}{u^r}\, du.$$

We can integrate the right-hand side by parts (with $f = \frac{1}{u^r}$ and $g' = e^{-u}$) to obtain

$$\Gamma(1-r) = \frac{1-e^{-u}}{u^r} \,|_{u=0}^{u=\infty} + r \int_0^\infty \frac{1-e^{-u}}{u^{r+1}}\, du.$$

Then consider the change of variables $s = \frac{u}{q}$ for $q > 0$

$$\frac{1}{r}\Gamma(1-r) = \int_0^\infty \frac{1-e^{-sq}}{(sq)^{r+1}} q\, ds = \frac{1}{q^r}\int_0^\infty \frac{1-e^{sq}}{s^{r+1}}\, ds$$

which gives

$$q^r = \frac{r}{\Gamma(1-r)} \int_0^\infty \frac{1-e^{-zq}}{z^{r+1}}\, dz$$

and we get our result by setting $q = X$ and using the Fubini-Tonelli Theorem.

For (43) let $s \geq 0$ and $z > 0$. By the definition of the Gamma function and the substitution $u = (\frac{t}{s})^z$ (where s and z are constants) we get

$$\Gamma(1/z) = \int_0^\infty u^{1/z-1} e^{-u}\, du = \int_0^\infty (\frac{t}{s})^{1-z} e^{-(t/s)^z} (\frac{t}{s})^{z-1} \frac{z}{s}\, dt$$

which leads to

$$s = \frac{z}{\Gamma(1/z)} \int_0^\infty e^{-(t/s)^z}\, dt.$$

Now we set $s = X^{-p}$ and $z = 1/p$. By integrability of X we can use Fubini-Tonelli, which yields our desired result. $\square$

The next proposition enables us to restrict ourselves to stock price processes with the property:

$$\mathbb{E}[[X]_T] < m$$

for some m $\in \mathbb{R}$. This makes the following proofs much easier.

We will need some basic results from probability theory which we summarize in the next lemma.

Lemma 2.5.

(A) Let $(U_n)_{n\geq 0}$ be a sequence of random variables, s.t. $\{U_n : n \geq 0\}$ is uniformly integrable. If we have $U_n \to U$ almost surely (a.s.) then we also have $U_n \to U$ in L_1.

(B) If $(U_n)_{n\geq 0}$ is a sequence of random variables, s.t. $U_n \in L_p$ and $U_n \to U$ in L_p then we have

$$\mathbb{E}[U_n|\mathcal{F}_t] \to \mathbb{E}[U|\mathcal{F}_t] \qquad \text{in } L_p \text{ as } n \to \infty.$$

(C) Let $(U_n)_{n\geq 0}$ and $(V_n)_{n\geq 0}$ be sequences of random variables, s.t. $U_n = V_n$ almost surely (a.s.) $\forall n \geq 0$. If U_n converges to a random variable U in L_1 and V_n converges to a random variable V a.s. . Then we have $U = V$ a.s..

Proof. See Hofbauer[15]. □

Proposition 2.6. *Let S be a price process with the assumptions of this section. Let h be a continuous function, for which we require that it is either increasing and nonnegative or that it is bounded.*
Let G be a measurable function, which has a decomposition $G = G_1 - G_2$, where G_1, G_2 are convex and $\mathbb{E}[G_{1,2}(S_T)] < \infty$.
If we have for all price processes $\tilde{S}_t$, which satisfy $[\tilde{X}]_T < m$ for some fixed $m \in \mathbb{R}$

$$\mathbb{E}[h([\tilde{X}]_T)|\mathcal{F}_t] = \mathbb{E}[G(\tilde{S}_T)|\mathcal{F}_t].$$

Then we have:

$$\mathbb{E}[h([X]_T)|\mathcal{F}_t] = \mathbb{E}[G(S_T)|\mathcal{F}_t]. \tag{44}$$

Proof. Let us define the process

$$\sigma_t^m := \sigma_t \mathbf{1}_{[X]_T \leq m}$$

and the process S_t^m by

$$dS_t^m := \sigma_t^m S_t^m dW_t$$

and finally set $X_t^m := \log(S_t^m)$.
Let us fix an ω and let $u \in \mathbb{R}$ s.t. $[X]_u(\omega) = m$ (if such an u does not exist we have $[X]_T(\omega) < m$). If $u \leq T$ we have $\int_0^T \sigma_t^{m^2}(\omega)\, dt = \int_0^u \sigma_t^2(\omega)\, dt = [X^m]_T(\omega) = [X]_u(\omega) = m$ and if $u > T$ we get $\int_0^T \sigma_t^{m^2}(\omega)\, dt = [X^m]_T(\omega) < m$.

Hence we have $\forall \omega$: $[X^m]_T \leq m$ and our assumption yields

$$\mathbb{E}[h([X^m]_T)|\mathcal{F}_t] = \mathbb{E}[G(S_T^m)|\mathcal{F}_t].$$

We have $[X^m]_T \to [X]_T$ as $m \to \infty$ a.s. .
By our assumption $\mathbb{E}[[X]_T] < \infty$ we have $\forall \omega \in N^c$ where $\mathbb{P}(N) = 0$ $[X]_T < \infty$. For a fixed $\omega \in N^c$ this implies that $[X]_T(\omega) < \tilde{m}$ where $\tilde{m} \in \mathbb{R}_+$. We see that $\forall m \geq \tilde{m}$ we have $[X^m]_T = [X]_T$. Therefore, $\forall \omega \in N^c$ we have $[X^m]_T \to [X]_T$ which proves the above statement.

We now distinguish the two possible assumptions for h.
We easily see that $[X^m]_T$ is increasing in m. Therefore, if h is increasing we have $h([X^m]_T)$ is increasing in m. By continuity of h we have $h([X^m]_T) \to h([X]_T$ almost surely. We have also assumed that $h \geq 0$. We can therefore apply the monotone convergence theorem for the conditional expectation to get:

$$\mathbb{E}[h([X^m]_T)|\mathcal{F}_t] \to \mathbb{E}[h([X]_T)|\mathcal{F}_t] \tag{45}$$

as $m \to \infty$ almost surely and in $\mathbf{L}^1(\mathbb{P})$.

On the other hand, if $h \leq K$ we use the dominated convergence theorem to get the same result.

Next, we will show that $\mathbb{E}[G(S_T^m)|\mathcal{F}_t] \to \mathbb{E}[G([S]_T)|\mathcal{F}_t]$ as $m \to \infty$ in in $\mathbf{L}^1(\mathbb{P})$.

There exist $\alpha, \beta \in \mathbb{R}$ and convex-nonnegative functions G_+, G_- with the following properties:
$G_\pm(S_0) = 0$ and $\mathbb{E}[G_\pm(S_T)] < \infty$ and

$$G(x) = G_+(x) - G_-(x) + \alpha x + \beta \qquad \forall x \in \mathbb{R}_+.$$

To see this, note that since G_1 is convex, there exists an affine function $g_1(x) = \alpha_1 x + \beta_1$ with $g \leq G_1$ s.t. $g_1(S_0) = G_1(S_0)$. Likewise, we find $g_2(x) = \alpha_2 x + \beta_2$ with $g_2(S_0) = G_2(S_0)$. We set $G_+(x) := G_1(x) - g_1(x)$ and $G_-(x) := G_2(x) - g_2(x)$, then we have the above properties since a sum of convex functions is again convex and $g_1 - g_2$ is of the form $\alpha x + \beta$.

We will show that $\mathbb{E}[G_+(S_T^m)|\mathcal{F}_t] \to \mathbb{E}[G_+(S_T)|\mathcal{F}_t]$ as $m \to \infty$ the rest works out completely analogous.

We have that $S_T^m \to S_T$ a.s. as $m \to \infty$.
As above, we have for a fixed $\omega \in N^c$ where N is a null set $[X]_t < \tilde{m}$. Therefore, we get $\forall m \geq \tilde{m}$ $\sigma_t^m = \sigma_t$, by the definition of S_t^m this implies $S_t^m = S_t$.

By convexity we have that G_+ is continuous, which yields that

$$G_+(S_T^m) \to G_+(S_T) \qquad \text{a.s. as } m \to \infty.$$

To proceed, we need to show that the family $G_+(S_T^m) : m \geq 0$ is uniformly integrable.
By assumption, we have $\mathbb{E}[G_+(S_T)] < \infty$. Therefore it suffices to show that:

$$\mathbb{E}[G_+(S_T^m)\mathbf{1}_{(G_+(S_T^m)>A)}] \leq \mathbb{E}[G_+(S_T)\mathbf{1}_{(G_+(S_T^m)>A)}] \qquad \forall m, A \geq 0.$$

We know that $G_+(x) \geq 0$, $G_+(S_0) = 0$ and that G_+ is convex. Combining these facts we see that there exist $a, b \in \mathbb{R}_+$ s.t

$$\mathbf{1}_{(G_+(x)>A)} = \mathbf{1}_{(x<S_0-a)} + \mathbf{1}_{(x>S_0+b)}.$$

Being the sum of convex functions the following function is convex

$$U(x) := G_+(x)\mathbf{1}_{(G_+(x)>A)} - \frac{A}{b}(x - S_0)\mathbf{1}_{(x>S_0+b)} - \frac{A}{b}(S_0 - x)\mathbf{1}_{(x<S_0-a)}.$$

In the following let $\mathcal{F}_T^\sigma \subset \mathcal{F}_t$ denote the σ-algebra generated by σ_t.
We then have

$$\begin{aligned}
&\mathbb{E}[G_+(S_T^m)\mathbf{1}_{(G_+(S_T^m)>A)}] \\
&= \mathbb{E}[U(S_T^m) + \frac{A}{b}(S_T^m - S_0)\mathbf{1}_{(S_T^m>S_0+b)} + \frac{A}{b}(S_0 - S_T^m)\mathbf{1}_{(S_T^m<S_0-a)}] \\
&= \mathbb{E}[\mathbb{E}[U(S_T^m) + \frac{A}{b}(S_T^m - S_0)\mathbf{1}_{(S_T^m>S_0+b)} + \frac{A}{b}(S_0 - S_T^m)\mathbf{1}_{(S_T^m<S_0-a)}|\mathcal{F}_T^\sigma]]
\end{aligned}$$

$$\leq \mathbb{E}[\mathbb{E}[U(S_T) + \frac{A}{b}(S_T - S_0)\mathbf{1}_{(S_T > S_0 + b)} + \frac{A}{b}(S_0 - S_T)\mathbf{1}_{(S_T < S_0 - a)}|\mathcal{F}_T^\sigma]]$$
$$= \mathbb{E}[G_+(S_T)\mathbf{1}_{G_+(S_T > A)}].$$

For the second inequality, we used the iterated conditioning property of the conditional expectation.

For the third inequality, we used that for a log-normal distributed random variable X with mean S_0 and $Var(\log(X)) = \sigma^2$. We observe that the formulas

$$\mathbb{E}[(S_0 - X)\mathbf{1}_{(X < S_0 - a)}] \text{ and } \mathbb{E}[(X - S_0)\mathbf{1}_{(X > S_0 + b)}]$$

are increasing in σ since the negative part is cut off and therefore more variance increases the value.
Furthermore $\mathbb{E}[U(X)]$ is increasing in σ by Jensen's inequality and convexity.

We have shown that $\{G_+(S_T^m) : m \geq 0\}$ is uniformly integrable.

By Lemma 2.5(A) this assures us together with $G_+(S_T^m) \to G_+(S_T)$ a.s. as $m \to \infty$ the convergence in L_1. From the convergence in L_1, it follows by Lemma 2.5(B) that

$$\mathbb{E}[G_+(S_T^m)|\mathcal{F}_t] \to \mathbb{E}[G_+(S_T)|\mathcal{F}_t] \qquad \text{as } m \to \infty \text{ in } L_1.$$

Finally we use our assumption $\mathbb{E}[h([X^m]_T)|\mathcal{F}_t] = \mathbb{E}[G(S_T^m)|\mathcal{F}_t]$ a.s. for $m \in \mathbb{N}$ and Lemma 2.5(C) with (45) to conclude

$$\mathbb{E}[h([X]_T)|\mathcal{F}_t] = \mathbb{E}[G(S_T)|\mathcal{F}_t] \qquad \text{a.s..}$$

□

The next proposition will be of great use in pricing the volatility swap later on. In a way the exponential functions of this proposition serve as basis functions with which we can get the price of financially interesting payoffs such as the volatility swap.

Proposition 2.7. *Assume that S_t satisfies the assumptions of this section. We have for $\lambda \in \mathbb{C}$ and $t \leq T$:*

$$\mathbb{E}[e^{\lambda[X]_T}|\mathcal{F}_t] = e^{\lambda[X]_t}\mathbb{E}[(S_T/S_t)^{1/2 \pm \sqrt{1/4 + 2\lambda}}|\mathcal{F}_t]. \tag{46}$$

Proof. By independence W_t is still a Brownian motion conditional on $\mathcal{F}_T^\sigma$,which is the σ-algebra generated by our volatility process. Therefore, we have conditional on $\mathcal{F}_t \vee \mathcal{F}_T^\sigma$:

$$X_T - X_t = \int_t^T \sigma_u \, dW_u - 1/2([X]_T - [X]_t).$$

The integrand of the Itô integral in the above expression is non-random and $\mathcal{F}_t$ measurable, therefore we know that the integral has a $N(0, \int_t^T \sigma_u^2 \, du)$ distribution. Under our continuity assumption, we know that $\int_t^T \sigma_u^2 \, du =$

$[X]_T - [X]_t$. Conditional on $\mathcal{F}_T^\sigma$, therefore $[X]_T - [X]_t$ is non random. We conclude that:

$$\int_t^T \sigma_u \, dW_u - 1/2([X]_T - [X]_t) \sim N(([X]_t - [X]_T)/2, [X]_T - [X]_t) \tag{47}$$

by the distribution of the integral. Let $p \in \mathbb{C}$:

$$\mathbb{E}[e^{p(X_T - X_t)}|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[e^{p(X_T - X_t)}|\mathcal{F}_t \vee \mathcal{F}_T^\sigma]|\mathcal{F}_t]$$

$$= \mathbb{E}[e^{\mathbb{E}[pX_T - pX_t|\mathcal{F}_t \vee \mathcal{F}_T^\sigma] + Var[pX_T - pX_t|\mathcal{F}_t \vee \mathcal{F}_T^\sigma]/2}|\mathcal{F}_t]$$

$$= \mathbb{E}[e^{\frac{(p^2-p)([X]_T - [X]_t)}{2}}|\mathcal{F}_t] = \mathbb{E}[e^{\lambda([X]_T - [X]_t)}|\mathcal{F}_t]$$

where we have set $\lambda = p^2/2 - p/2$.
We have the first equality by the iterated-conditioning property of the conditional expectation. The second equality, follows from the well known form of the characteristic function of the normal distribution. The third equality follows from (47). □

Remark As a sanity check one can differentiate the above formula with respect to p and evaluate at $p = 0$ to obtain

$$\mathbb{E}[[X]_T] = -2\mathbb{E}[X_T]$$

which is the familiar result we derived in the variance swap section.

We will now find formulas to price volatility derivatives in a general form.

Remark The following formulas might look somewhat unusual at first glance. In fact one could derive simpler formulas under our assumptions. To derive these formulas one could almost copy our proofs and would not need the $\alpha_\pm$ terms. However, our formulas have the advantage of being correlation neutral in the sense of Carr-Lee (see Carr and Lee [6] Remark 4.3 and Definition 4.4). A formula which is correlation neutral has the property that for ρ near zero we have an error of only $O(\rho^2)$. Later we wish to test how well behaved our formulas are when the stock-process has jumps. Therefore we wish to derive the most practical and useful formulas possible.

For all our pricing formulas we assume that the stock price satisfies the assumptions of this section. The proof of the next theorem generalizes the one of Carr and Lee [6], where the special case of $r = 1/2$ is proven.

Theorem 2.8. *We have for* $0 < r < 1$*:*

$$\mathbb{E}[[X]_T{}^r|\mathcal{F}_t] = \mathbb{E}[f_r(S_T, S_t, [X]_t)|\mathcal{F}_t] \tag{48}$$

where

(49)

$$f_r(S, u, q) := \frac{r}{\Gamma(1-r)} \int_0^\infty \alpha_+(z) \frac{1 - e^{-zq}(S/u)^{p_+}}{z^{r+1}} + \alpha_-(z) \frac{1 - e^{-zq}(S/u)^{p_-}}{z^{r+1}} \, dz$$

and we set
$\alpha_{\pm}(z) := 1/2 \mp \frac{1}{2\sqrt{1-8z}}$ *and* $p_{\pm}(z) := 1/2 \pm 1/2\sqrt{1-8z}$. *Moreover,* f_r *is integrable and convergent.*

Proof. By Proposition 2.6 it suffices to prove the formula for a price process S with $[X]_T < m$ with $m \in \mathbb{R}$. Let q be $\mathcal{F}_t$ measurable. We use (42) to get

$$\mathbb{E}[\sqrt{[X]_T - [X]_t - q}|\mathcal{F}_t] = \frac{r}{\Gamma(1-r)}\mathbb{E}[\int_0^\infty \frac{1 - e^{-z([X]_T - [X]_t - q)}}{z^{r+1}}\, dz|\mathcal{F}_t].$$

Next we want to apply the Fubini Tonelli Theorem, which can be applied since we have $|1 - e^{-z(X]_T - [X]_t - q)}| < 1 - e^{-z(m+q)}$. We use the series representation of the exponential and see that the integral is finite.
We have:

$$\begin{aligned}
&\frac{r}{\Gamma(1-r)}\mathbb{E}[\int_0^\infty \frac{1 - e^{-z([X]_T - [X]_t - q)}}{z^{r+1}}\, dz|\mathcal{F}_t] \\
&= \frac{r}{\Gamma(1-r)}\int_0^\infty (\alpha_+(z) + \alpha_-(z))\frac{1 - \mathbb{E}[e^{-z([X]_T - [X]_t - q)}|\mathcal{F}_t]}{z^{r+1}}\, dz \\
&= \frac{r}{\Gamma(1-r)}\int_0^\infty \alpha_+(z)\frac{1 - e^{-zq}\mathbb{E}[e^{p_+(X_T - X_t)}|\mathcal{F}_t]}{z^{r+1}} + \alpha_-(z)\frac{1 - e^{-zq}\mathbb{E}[e^{p_-(X_T - X_t)}|\mathcal{F}_t]}{z^{r+1}}\, dz
\end{aligned}$$

where we used Fubini Tonelli and $\alpha_+(z) + \alpha_-(z) = 1$ for the first equality and (46) for the second one.
In order to get our result we have to apply Fubini's Theorem once more. We will justify this in the following.

$$\begin{aligned}
A_{\pm}(z) &:= \mathbb{E}[|1 - e^{-qz + (1/2 \pm \sqrt{1/4 - 2z})(X_T - X_t)}|^2|\mathcal{F}_t] \\
&\le \mathbb{E}[|1 - e^{-qz + (1/2 \pm \sqrt{1/4 - 2z})(X_T - X_t)}||\mathcal{F}_t]^2
\end{aligned}$$

by Jensen's inequality. We have $\mathbb{E}[|1 - e^{-qz + (1/2 \pm \sqrt{1/4 - 2z})(X_T - X_t)}||\mathcal{F}_t]^2 \to 1$ as z $\to \infty$, therefore:
(*)

$$\mathbb{E}[\frac{|\alpha_{\pm}(z) + 1 - e^{-qz + (1/2 \pm \sqrt{1/4 - 2z})(X_T - X_t)}|}{z^{r+1}}|\mathcal{F}_t] = O(z^{-r-1}) \quad \text{as } z \to \infty.$$

Next, we have for $z < 1/8$ that term in the absolute values in the above formula is real, therefore we calculate

$$\begin{aligned}
A_{\pm}(z) &\le \mathbb{E}[1 - 2e^{-qz + (1/2 \pm \sqrt{1/4 - 2z})(X_T - X_t)} + e^{-2qz + 2(1/2 \pm \sqrt{1/4 - 2z})(X_T - X_t))}|\mathcal{F}_t] \\
&= 1 - 2e^{-qz}\mathbb{E}[e^{-z([X]_T - [X]_t)}|\mathcal{F}_t] + e^{-2qz}\mathbb{E}[e^{\frac{1 - 8z \pm \sqrt{1-8z}}{2}([X]_T - [X]_t)}|\mathcal{F}_t]
\end{aligned}$$

where we used (46) in the second step.
For the moment generating function we have $v(\xi) := \mathbb{E}[e^{\xi[X]_T}|\mathcal{F}_t] < e^{m\xi}$.Therefore v is analytic.

As z $\to 0$ we immediately observe

$$A_+(z) = O(1)$$

and using the analyticity

$$A_-(z) = 1 - 2(2 - zm'(0) - qz + O(z^2)) + 1 - 2zm'(0) - 2qz + O(z^2) = O(z^2)$$

where we used $(\frac{1-8z\pm\sqrt{1-8z}}{2})'\Big|_{z=0} = -2$.

We observe that $\alpha_+(z) = O(z)$ and $\alpha_-(z) = O(1)$ when we let $z \to 0$ and get

(**)

$$\mathbb{E}[\frac{|\alpha_+(z)[1-e^{-qz+(1/2+\sqrt{1/4-2z})(X_T-X_t)}]|}{z^{r+1}}|\mathcal{F}_t] = \frac{|\alpha_+(z)|\sqrt{A_+(z)}}{z^{r+1}} = \frac{O(z)O(1)}{z^{r+1}} = O(z^{-r})$$

and

(***)

$$\mathbb{E}[\frac{|\alpha_-(z)[1-e^{-qz+(1/2-\sqrt{1/4-2z})(X_T-X_t)}]|}{z^{r+1}}|\mathcal{F}_t] = \frac{|\alpha_-(z)|\sqrt{A_-(z)}}{z^{r+1}} = \frac{O(1)O(z)}{z^{r+1}} = O(z^{-r})$$

as $z \to 0$.

We use this to see that for any $u \in \mathbb{R}$

$$\begin{aligned}
&\frac{r}{\Gamma(1-r)}(\int_0^\infty \alpha_+(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_+(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}} + \alpha_-(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_-(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}}\,dz) \\
&= \frac{r}{\Gamma(1-r)}(\int_0^u \alpha_+(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_+(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}}\,dz + \int_0^u \alpha_-(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_-(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}}\,dz \\
&+ \int_u^\infty \alpha_+(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_+(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}}\,dz + \int_u^\infty \alpha_-(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_-(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}}\,dz).
\end{aligned}$$

The first integral is finite by (**) and the fact that $\int_0^u z^{-r}\,dz < \infty$. The second one by (***) and the last two by (*). Therefore, we have shown that we can use Fubini Tonelli to get

$$\begin{aligned}
&\frac{r}{\Gamma(1-r)}\int_0^\infty \alpha_+(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_+(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}} + \alpha_-(z)\frac{1-e^{-zq}\mathbb{E}[e^{p_-(X_T-X_t)}|\mathcal{F}_t]}{z^{r+1}}\,dz \\
&\quad= \frac{r}{\Gamma(1-r)}\mathbb{E}[\int_0^\infty \alpha_+(z)\frac{1-e^{-zq}e^{p_+(X_T-X_t)}}{z^{r+1}} + \alpha_-(z)\frac{1-e^{-zq}e^{p_-(X_T-X_t)}}{z^{r+1}}\,dz|\mathcal{F}_t].
\end{aligned}$$

Setting $q := [X]_t$ gives the result. □

We use this theorem directly to price the volatility swap. In finance the volatility swap is the most important derivative of this section, so we will therefore make a corollary out of it.

Corollary 2.9. *Under the assumptions of this section we have:*

$$\mathbb{E}[\sqrt{[X]_T}|\mathcal{F}_t] = \mathbb{E}[f(S_T, S_t, [X]_t)|\mathcal{F}_t] \tag{50}$$

where

$$f(S,u,q) := \frac{1}{2\sqrt{\pi}} \int_0^\infty \alpha_+(z) \frac{1 - e^{-zq}(S/u)^{p_+}}{z^{3/2}} + \alpha_-(z) \frac{1 - e^{-zq}(S/u)^{p_-}}{z^{3/2}} \, dz \tag{51}$$

with $\alpha_\pm(z)$ and $p_\pm$ from Theorem 2.8. Moreover, f is integrable and convergent.

Proof. Follows directly by Theorem 2.8, by setting $r = \frac{1}{2}$ and using $\frac{r}{\Gamma(1-r)} = \frac{1}{2\sqrt{\pi}}$. □

The next corollary provides a practical and easy formula. The Bessel function is implemented in most mathematical software packages such as Matlab.

Definition 2.10. *Let $\nu \geq -1$ and $x > 0$, then the modified Bessel function I_ν is defined by*

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{2k+\nu}}{k!} \Gamma(\nu + k + 1).$$

Corollary 2.11. *At inception time we have for the price volatility swaps*

$$\mathbb{E}[\sqrt{[X]_T}] = \mathbb{E}[\sqrt{\frac{\pi}{2}} |yI_0(\frac{y}{2}) - yI_1(\frac{y}{2})|] \tag{52}$$

where $y := \log(S_T/S_0)$ and $I_{0,1}$ is the modified Bessel function of first resp. second order.

Proof. Follows by integration of (50), see Carr and Lee [6] (Corollary 6.11). □

The next corollary gives the price of the volatility swap in terms of European call and put options. Thus, we relate, as in the variance swap case, the value of the volatility swap to option prices. These option prices can be obtained from the market.

Corollary 2.12. *Let $C(K), P(K)$ denote call/put prices at strike K. At time $t = 0$ we have:*

$$\begin{aligned}
\mathbb{E}[\sqrt{[X]_T}] =& \frac{\sqrt{\pi/2}}{S_0}(C(S_0) + P(S_0)) \\
&+ \int_{S_0}^\infty \sqrt{\frac{\pi}{8K^3 S_0}} (I_1(\log(\sqrt{K/S_0})) - I_0(\log(\sqrt{K/S_0}))) C(K) dK \\
&+ \int_0^{S_0} \sqrt{\frac{\pi}{8K^3 S_0}} (I_0(\log(\sqrt{K/S_0})) - I_1(\log(\sqrt{K/S_0}))) P(K) dK
\end{aligned}$$

and at a general time $t \leq T$ we have:

$$\begin{aligned}\mathbb{E}[\sqrt{[X]_T}|\mathcal{F}_t] =& \sqrt{[X]_t}\\ &+ \frac{1}{\sqrt{\pi}}\int_{S_t}^{\infty}[\int_0^{\infty}\frac{e^{-z[X]_t}}{K^2\sqrt{z}}\alpha_+(z)(K/S_t)^{p_+} + \alpha_-(z)(K/S_t)^{p_-}\,dz]C(K)\,dK\\ &+ \frac{1}{\sqrt{\pi}}\int_0^{S_t}[\int_0^{\infty}\frac{e^{-z[X]_t}}{K^2\sqrt{z}}\alpha_+(z)(K/S_t)^{p_+} + \alpha_-(z)(K/S_t)^{p_-}\,dz]P(K)\,dK\end{aligned}$$

with $\alpha_\pm$ and $p_\pm$ as before.

Proof. Apply Theorem 1.12 to the above formulas. □

As a by-product we can derive the price of some more volatility derivatives. The following derivatives are called "inverse variance" claims.

Proposition 2.13. *Let $r, \epsilon > 0$, we then have*

$$\mathbb{E}[([X]_T + \epsilon)^{-r}|\mathcal{F}_t] = \mathbb{E}[f_{-r}(S_T, S_t, [X]_t + \epsilon)|\mathcal{F}_t] \tag{53}$$

where we set

$$f_{-r}(S, u, q) := \frac{1}{r\Gamma(r)}\int_0^{\infty}(\alpha_+(z)(S/u)^{p_+} + \alpha_-(z)(S/u)^{p_-})e^{-z^{1/r}}\,dz$$

and $\alpha_\pm(z) := 1/2 \mp \frac{1}{2\sqrt{1-8z^{1/r}}}$, $p_\pm(z) := 1/2 \pm \sqrt{1/4 - 2z^{1/r}}$.

Proof. As usual by Proposition 2.6 we can restrict ourselves to the case when $[[X]_T] < m$. We get by (43)

$$\mathbb{E}[([X]_T + \epsilon)^{-r}|\mathcal{F}_t] = \frac{1}{\Gamma(r)r}\mathbb{E}[\int_0^{\infty} e^{-z^{1/r}([X]_T+\epsilon)}\,dz|\mathcal{F}_t].$$

We have $e^{-z^{1/r}([X_T]+\epsilon)} < e^{-z^{1/r}(m+\epsilon)}$ which suffices to apply Fubini-Tonelli and we get

$$\begin{aligned}&\frac{1}{\Gamma(r)r}\mathbb{E}[\int_0^{\infty} e^{-z^{1/r}([X]_T+\epsilon)}|\mathcal{F}_t]\,dz\\ &= \frac{1}{\Gamma(r)r}\int_0^{\infty} e^{-z^{1/r}([X]_t+\epsilon)}\mathbb{E}[e^{-z^{1/r}([X]_T-[X]_t)}|\mathcal{F}_t]\,dz\\ &= \frac{1}{\Gamma(r)r}\mathbb{E}[\int_0^{\infty}(\alpha_+(z)e^{p_+(X_T-X_t)} + \alpha_-(z)e^{p_-(X_T-X_t)})e^{-z^{1/r}([X]_t+\epsilon)}\,dz|\mathcal{F}_t].\end{aligned}$$

For the second equality we have used $\alpha_+(z) + \alpha_-(z) = 1$, (46) and Fubini-Tonelli once more, which is justified by

$$|e^{1/2\pm\sqrt{1/4-2z^{1/r}}(X_T-X_t)}| \leq 1.$$

□

2.4. **The impact of jumps, numerical tests.** In this section we want to test the formulas for the fair value of the volatility swap proven above when jumps are present. We do this by underlying specific models.

Of course, when we have a model in which we really believe we can price our volatility swap with Monte Carlo simulation. However, in practice we do not know which process volatility follows. Our theory is therefore of great value. Nevertheless, we still have the problem that we assume continuity, while it is known that to generate the observed volatility surface an additional jump process in our model is usually needed. Thus, we would like to know, if our theory still works well enough when the price process has jumps of reasonable size and frequency. In the following tests, we get the true value of the volatility swap in the different models with Monte Carlo simulation.

We start by introducing the Bates model, which is a model of SVJ class (stochastic volatility with jumps). Note that of course this model for the price process does not satisfy our continuity assumption which was made in this section. The stock price follows the following dynamics:

$$dS_t = \sqrt{v_t}S_t dW_t^1 + (e^{\alpha+\delta\epsilon} - 1)S_t dq \tag{54}$$

where $\alpha, \delta \in \mathbb{R}$ and ϵ is $N(0,1)$ distributed and q is a Poisson process. We have a Heston-style volatility process:

$$dv_t = \kappa(\theta - v_t)dt + \xi v_t W_t^2 \tag{55}$$

where $\lambda, \theta, \xi \in \mathbb{R}_+, \kappa \in \mathbb{R}$ and W_t^1, W_t^2 are Brownian motions.

We let $[dW^1, dW^2] = 0$, that means our Brownian motions are independent.

For the Poisson process q we intuitively have:

$$dq = 0 \text{ with probability } (1-\lambda)dt \text{ and } dq = 1 \text{ with probability } \lambda dt. \tag{56}$$

In our model we have the following parameters:

α	controls the mean value of our jump
δ	controls the variance of our jumps
λ	controls how often jumps occur
θ	gives the mean value of our volatility process
ξ	the volatility of the volatility
κ	controls how tightly the volatility is bound to θ

Remark In practice we will typically choose α to be negative, because jumps are going downward most of the time. If we choose $\lambda = 0$ or $\alpha, \beta = 0$, we obtain the well known Heston model which of course satisfies our assumptions.

Remark One could make the above model more realistic by introducing a correlation between the driving Brownian motion of the stock price and that of the volatility process. This correlation would typically be negative.

However, we have chosen not to do this. Therefore, the errors that we will get in the following tests below derive from the fact that the models do not satisfy the continuity assumption only.

This table gives the parameters we used for the graphs involving the Bates model.

Figure	λ	α	δ	θ	ξ	κ	T
1	0.3	-0.3	0.2	0.04	0.39	1.05	3
2	0 to 1	-0.3	-0.2	0.04	0.39	1.05	2
3	0.3	-1 to 1	0.2	0.04	0.39	1.05	2
4	0.3	0	0 to 1	0.04	0.39	1.05	2

In Figure 1 we have a sample path of our process. We see that in the case of this particular path, there is a jump downward at approximately $t = 2$.
In Figure 2 we see that the true value and the Bessel formula coincide at zero, because it is at this point that our assumptions are satisfied. The true value becomes larger, because of the additional volatility from the jumps, which become more and more frequent when λ gets larger. The Bessel formula is more sensitive to changes in λ and overestimates the value of the volatility swap more and more.
In Figure 3 we vary the average jump size. Obviously there is almost no error at a jump size near zero. The Bessel formula always overestimates the fair value. We get a much smaller error when α is negative compared to when α is positive.
Figure 4 shows that varying δ has no real impact on the error. Of course the value goes up when δ increases, because then we have jumps. For small δ the values nearly coincide, because there the jumps play no role.(when we choose $\alpha = 0$).
All these characteristics stay the same for other reasonable choices of the other parameters. We provide a table where we test the formula with the same jump parameters with which we tested the variance swap in the previous section. The parameters for the volatility process are taken as in the graphs. The error is given in terms of the true value of the volatility swap.

Author(s)	λ	α	δ	error
Bakshi, Cao and Chen[2]	0.59	-0.05	0.07	0.70%
Matytsin [3]	0.5	-0.15	0	0.95%
Duffie, Pan, and Singleton[11]	0.11	-0.12	0.15	1.06%
Jim Gatheral[13]	0.13	-0.12	0.10	1.08%

Although the errors are slightly larger than the ones we got for the variance swap, the errors are still very small.

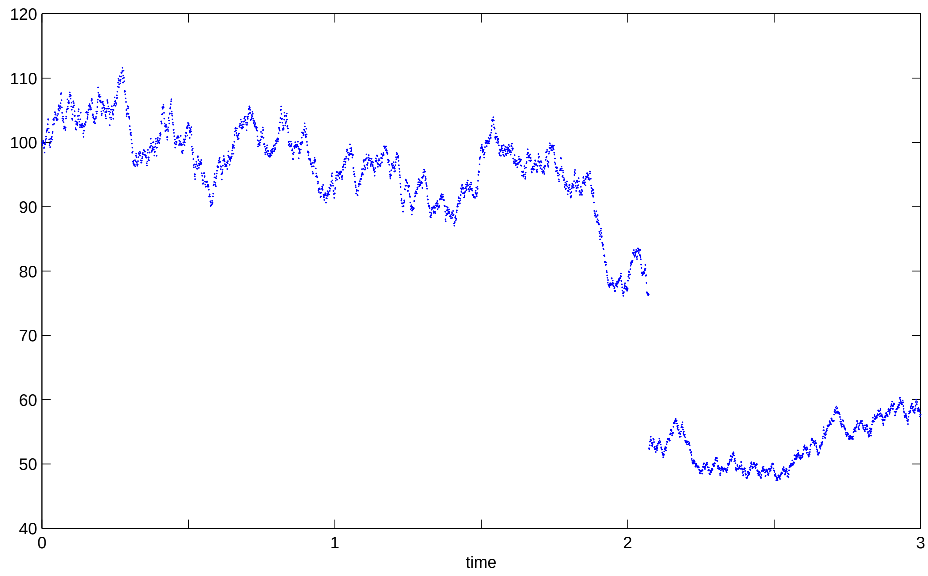

FIGURE 1. sample path-Bates model

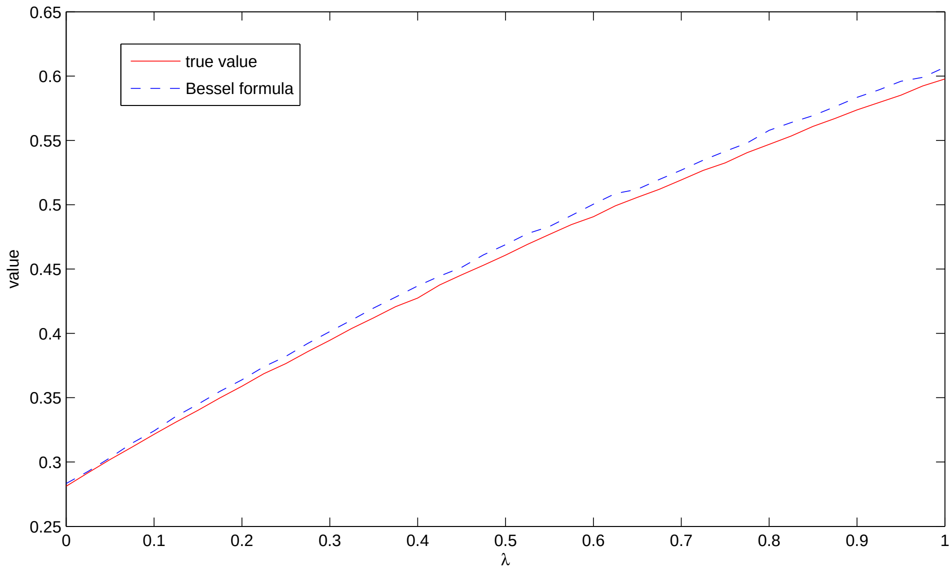

FIGURE 2. Bates model-Bessel formula vs. true value plotted against λ

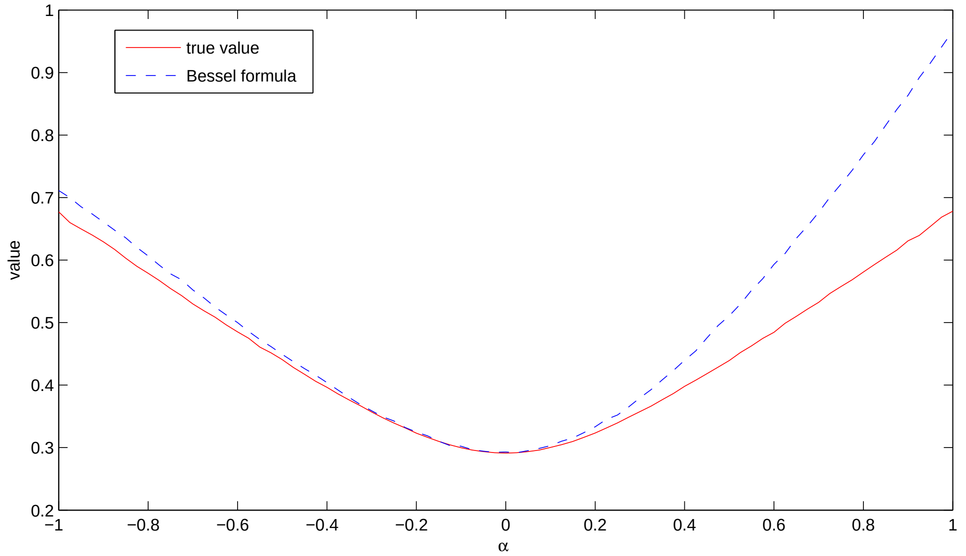

FIGURE 3. Bates model-Bessel formula vs. true value plotted against α

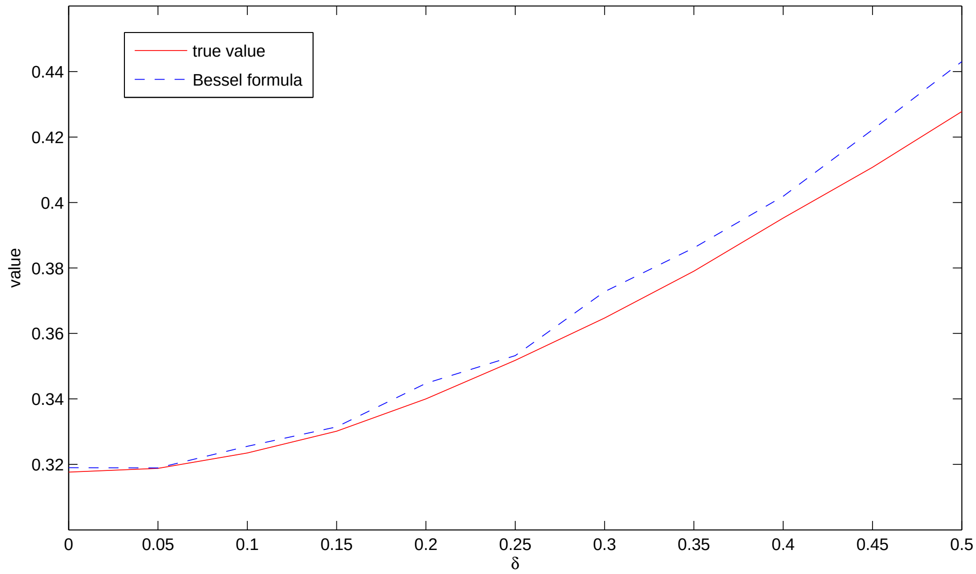

FIGURE 4. Bates model-Bessel formula vs. true value plotted against δ

Next, we test our formula with another process with jumps. The NIG process X_t^{NIG} is a Lévy process, i.e. a process which has independent and stationary increments. In the case of the NIG process X_t has a $NIG(\alpha, \beta, t\delta)$ law. For the parameters we require $\alpha > 0, -\alpha < \beta < \alpha, \delta > 0$
The NIG law is defined via it is characteristic function which is given by:

$$\phi_{NIG}(u, \alpha, \beta, \delta) = e^{-\delta(\sqrt{\alpha^2-(\beta+iu)^2}-\sqrt{\alpha^2-\beta^2})}$$

The NIG process can be related to an Inverse Gaussian time-changed Brownian Motion. The Inverse Gaussian Process is a Lévy process which has at time t the $IG(at, b)$ distribution. The density function of the $IG(a, b)$ distribution is for $x > 0$ given by

$$f_{IG}(x) = \frac{a}{\sqrt{2\pi}} e^{ab} x^{-3/2} e^{-1/2(a^2x^{-1}+b^2x)}.$$

Then our NIG process X_t can be written as

$$X_t^{NIG} = \beta\delta^2 I_t + \delta W_{I_t} \tag{57}$$

where I_t is an IG process with parameters $a = 1$ and $b = \delta\sqrt{\alpha^2 - \beta^2}$ and W_t is a Brownian motion.

The following characteristics can be calculated.

	$X_t^{NIG}(\alpha, \beta, \delta)$	$X_t^{NIG}(\alpha, 0, \delta)$
mean	$\frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}$	0
variance	$\frac{\alpha^2\delta}{(\alpha^2-\beta^2)^{-3/2}}$	$\frac{\delta}{\alpha}$
skewness	$\frac{3\beta\alpha^{-1}\delta^{-1/2}}{(\alpha^2-\beta^2)^{1/4}}$	0
kurtosis	$3(1+\frac{\alpha^2+4\beta^2}{\delta\alpha^2\sqrt{\alpha^2-\beta^2}})$	$3(1+\delta^{-1}\alpha^{-1})$

Our stock price process is then given by

$$S_t = e^{X_t^{NIG}+mt}$$

where m is an additional drift parameter which makes our process S_t a martingale. In our case assuming zero interest rates we have:

$$m = \delta(\sqrt{\alpha^2 - (\beta^2 + 1)^2} - \sqrt{\alpha^2 - \beta^2}).$$

This table provides the parameters we use for our tests:

Figure	α	β	δ
5	70	-40	0.5
6	40	-30 to 30	0.08
7	10 to 50	8	0.08

In Figure 5, we plot a sample path of our stock price process. We recognize the large downward jumps due to the high β that we have chosen. Therefore, our drift m has to be positive to make our process a martingale.

In Figure 6 we see that for all sizes of β the Bessel formula overestimates the true value. At $\beta = 0$ we have almost no error. Looking at the IG density function we find that the IG-process looks almost like a straight line for small β and large α. We see from (57) that the discontinuity almost vanishes numerically. For larger values of β the error gets larger and larger. At $\beta = -30$ where jumps become extreme, we have an error of 16%.

Figure 7 shows that for large α the Bessel formula is a very accurate approximation to the true value. In this Figure, the Bessel formula again always overestimates the true value. The error gets a larger for small values of α due to the higher probability of large jumps.

Remark In the variance swap case we had the rule of thumb that higher down jumps imply that our strategy (which holds true in the continuous case) underestimates the true value and vice versa. Figure 3 and Figure 6 make clear that for the volatility swap this is not the case.
For the processes we have studied so far, the Bessel formula overestimates the price when jump sizes increase in either direction.
From (23) we have seen that large β and small α increase the extent of the error in the variance swap case. This pattern prevails in our analysis here, which does not surprise us since choosing the parameters in this way makes discontinuities more extreme.

This table shows the error in terms of the true value for some calibrated parameters. We get the parameters from Rydberg [18]. (choosing $T = 2$ and $n = 100$)

α	β	δ	error
25.86	0.33	0.014	4.98%
31.66	0.008	0.0095	1.64%
75.49	-4.089	0.012	2.01%

We see that the errors are larger compared to the ones we got for the Bates model, but the approximation is still very good.

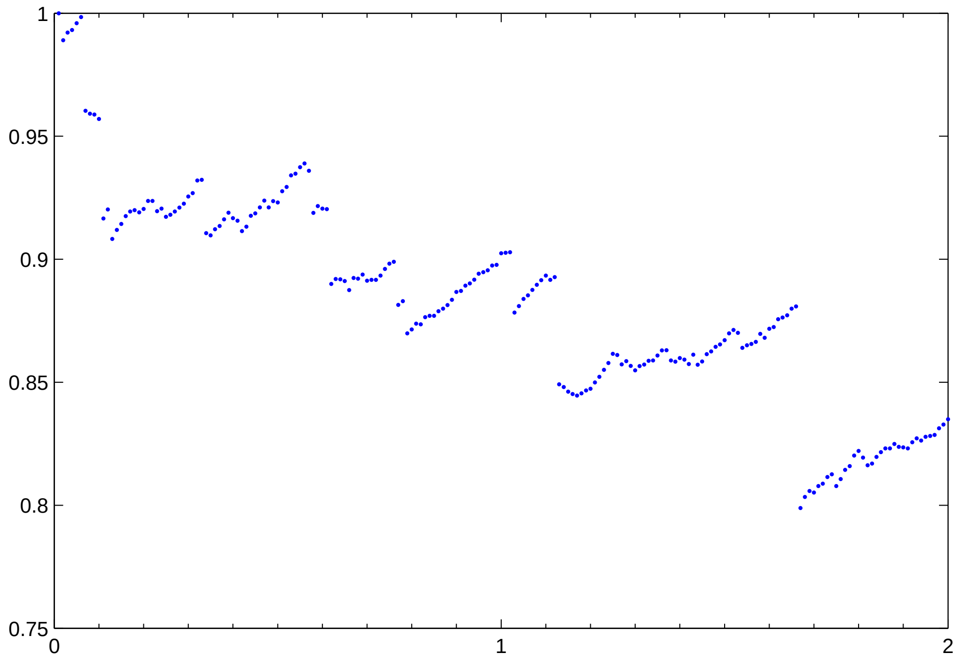

FIGURE 5. sample stock path-NIG process

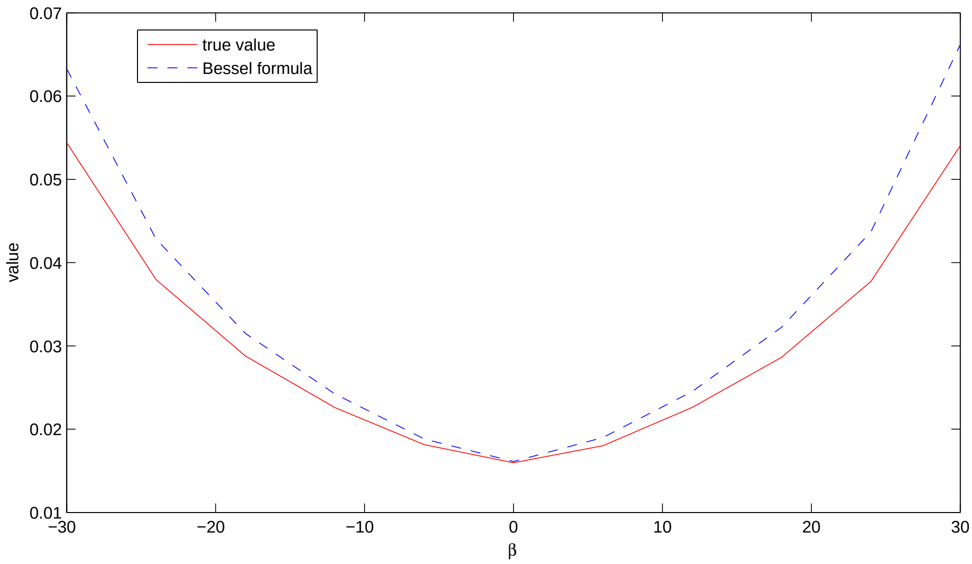

FIGURE 6. NIG model-Bessel formula vs. true value plotted against β

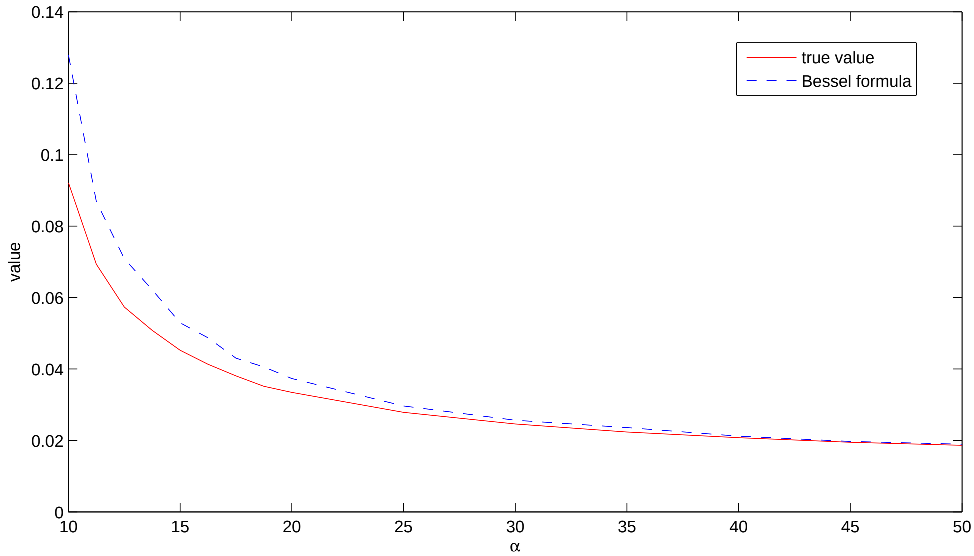

FIGURE 7. NIG model-Bessel formula vs. true value plotted against α

Finally we will test our formula using the VG process.
Our VG-process X_t^{VG} with parameters C, G, M is a Lévy process and can be written as

$$X_t^{VG} = G_t^1 - G_t^2 \tag{58}$$

where G_t^1 is a Gamma(C, G)-process and G_t^2 is a Gamma(C, M)-process.
A Gamma(a, b)-process is defined as the stochastic process starting at zero, which has stationary and independent Gamma(a, b) distributed increments. The Gamma(a, b) distribution has the density function defined for $x > 0$:

$$f_{Gamma}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-xb}.$$

The Gamma-process is increasing and (58) suggests why choosing $G > M$ results in higher up jumps.

We provide the characteristics of the VG-process:

	$X_t^{VG}(C, G, M)$	$X_t^{VG}(C, G, G)$
mean	$\frac{C(G-M)}{MG}$	0
variance	$\frac{C(G^2+M^2)}{(MG)^2}$	$2CG^{-2}$
skewness	$\frac{2C^{-1/2}(G^3-M^3)}{(G^2+M^2)^{3/2}}$	0
kurtosis	$\frac{3(1+2C^{-1}(G^4+M^4)}{(M^2+G^2)^2}$	$3(1 + C^{-1})$

Alternatively we can understand the VG-process with another representation. The VG-process can be written in terms of parameters θ, ν, σ as

$$X_t^{VG} = \theta G_t + \sigma W_{G_t}$$

where G_t is a Gamma(ν, ν) process.
The important property for us is that $\theta = 0$, if and only if $G = M$. Then the distribution is symmetric and is just a Gamma-time changed Brownian motion. If $G > M$ we have a negative θ which leads to negative skewness. In any case,we see that the assumptions of this section are not satisfied. Our process is not continuous.

Our stock price process is given by:

$$S_t = e^{X_t^{VG} + mt}.$$

Similar to the NIG case m is needed to make S_t a martingale, we have:

$$m = C \log((M-1)(G+1)/(MG)).$$

We proceed to test our formula when the underlying stock price process is modeled as above.

This table shows the parameters we used for our plots:

Figure	C	G	M	T
8	10	25	10	1
9	10	30	5 to 60	1
10	2.5 to 20	20	40	1

In Figure 9 we see that the pattern from the NIG case prevails in the VG-model. The Bessel formula always slightly overestimates the true value, but the approximation is very good. We see that at $M = 5$, where we have a very high volatility coming from the large upward jumps, the error becomes slightly larger. For high values of M we nearly have a perfect match of the values.
The next Figure shows that the parameter C which primarily controls the variance and the kurtosis has no real impact on the difference between the values. Of course, the value of the volatility swap goes up as C increases.

We continue our program with a table of errors in terms of the true value (setting $n = 100, T = 2$) with calibrated parameters from the literature:

Reference	C	G	M	error
Carr, Chang, Madan [4]	5.94	20.26	39.7	4.61%
Fiorani [12]	7.51	14.11	33.29	4.94%
Itkin [16]	6.25	14.4	60.24	6.28%

For the VG-process we get the largest errors so far, but we see that the Bessel formula is still a useful approximation.

Remark After having tested the Bessel formula in three different models with jumps, we have seen that it works remarkably well when the continuity assumption is not satisfied. For reasonable parameters the error always stays below 6.5 percent. Additionally we have seen that the Bessel formula always overestimated the true value.
Our tests suggest that the strategy can be useful as an upper bound, even for someone who strongly beliefs that a good model must contain jumps.
We have shown that we can use the formula to price the volatility swap via call/put options, in a liquid market we then of course do not even have to think about specifying any particular model.

Remark We have simulated all of the above processes using Matlab. For the Bates model we just have to discritize the SDE's and simulate a Poisson process. For the VG-process the representation as a difference of Gamma process is helpful. We simulated the NIG-process as a time-changed Brownian motion. In Schoutens [19](chapter 8) one can find instructions on how to simulate such paths and more details concerning the VG- and the NIG-process.

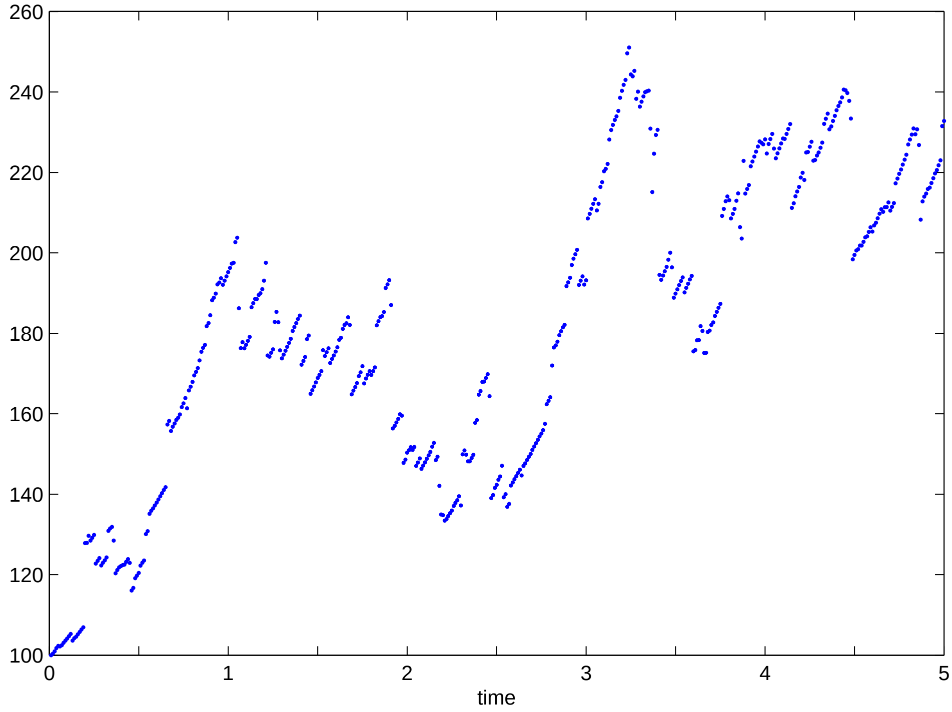

FIGURE 8. VG process-sample path

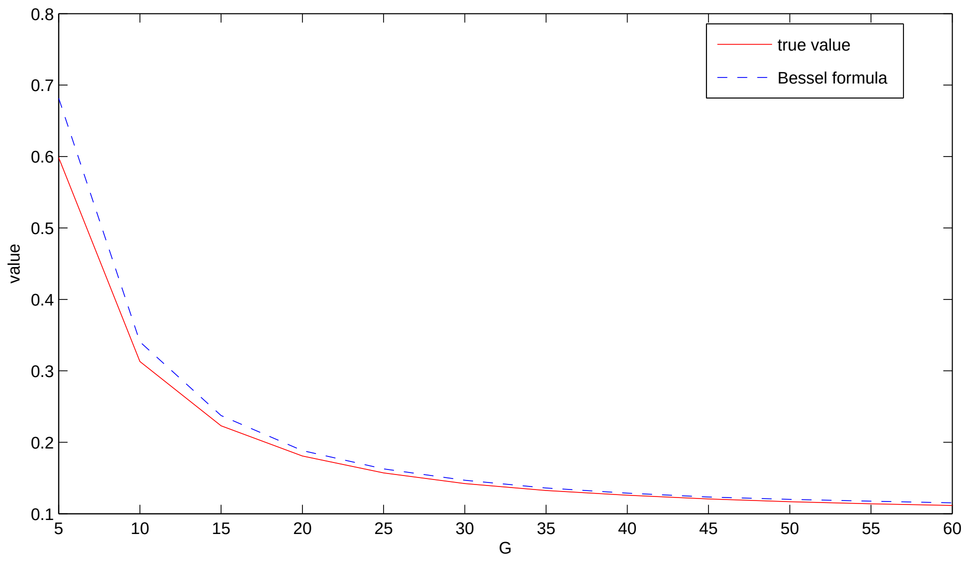

FIGURE 9. VG model-Bessel formula vs. true value plotted against M

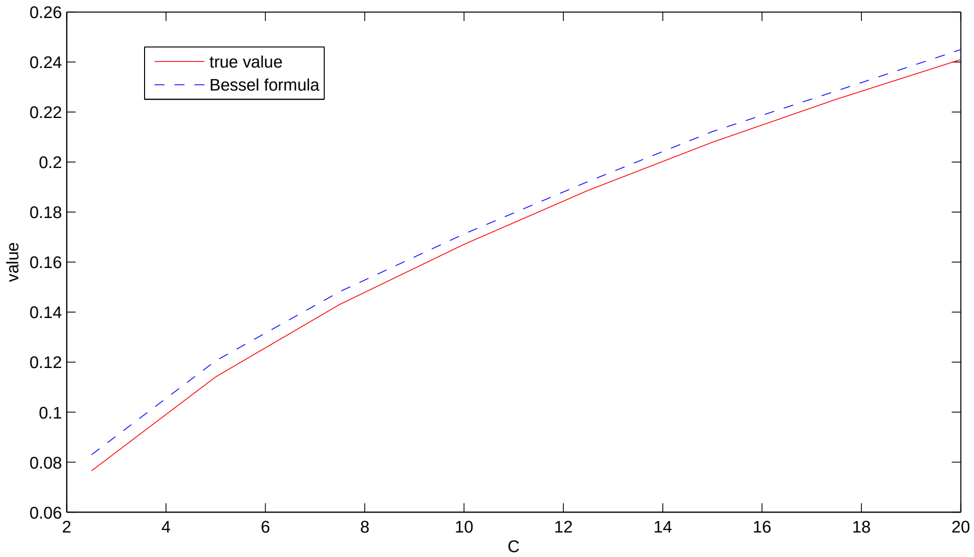

FIGURE 10. VG model-Bessel formula vs. true value plotted against C

3. Options on variance

3.1. **Assumptions.** In this section we work under the same very general assumptions as in the first section where we replicated the variance swap (see Section 1.1).

Without assuming independence, we do not find exact replication strategies or exact formulas, but we find sub- and superreplication strategies, which are of great practical importance.

3.2. **Hedging options on variance.** In this section we deal with a contract called the variance call. The variance call pays at some fixed time T and for some fixed Q the amount

$$([X]_T - Q)^+ \tag{59}$$

We will need the following two propositions:

Proposition 3.1. *Let U be an open set with $(S_0, 0) \in U \subset \mathbb{R}^2$. Let $g \in C^{2,1}$ on U and continuous on $\bar{U}$ and let $T \in (0, \infty)$. We then have for all stopping times with $\tau \leq \inf(t : (S_t, [\phi(S)]_t) \notin U)$*

$$g(S_{T\wedge\tau}, [X^\phi]_t) = g(S_0, 0) + \int_0^{T\wedge\tau} g_y d\, S_u + \int_0^{T\wedge\tau} (\frac{1}{2} g_{yy} + \phi_y^2 g_q)\, dS_u \tag{60}$$

and if $\phi_y(y) > 0, \forall y \in\ U$ then

$$g(S_{T\wedge\tau}, [X^\phi]_t) = g(S_0, 0) + \int_0^{T\wedge\tau} g_y d\, S_u + \int_0^{T\wedge\tau} (\frac{1 g_{yy}}{2\phi_y^2} + g_q)\, d[X^\phi]_u. \tag{61}$$

Proof. Define $\tau_n := \inf\{T : \exists (y, q) \in \mathbb{R} \times \mathbb{R}_+\ U s.t. |S_t - y| + |[X_t^\phi] - q| < 1/n\}$. We can apply Itô's rule to $(S_{t\wedge\tau_n}, [X^\phi]_{t\wedge\tau_n}) \forall n$ since the stopped process only takes values in U where g is in $C^{2,1}$ by assumption, we have $\forall\ T$

$$g(S_{T\wedge\tau_n}, [\phi(S)]_{T\wedge\tau_n}) = g(S_0, 0) + \int_0^{T\wedge\tau_n} g_y d\, S_u + \int_0^{T\wedge\tau_n} (\frac{1}{2} g_{yy} + \phi_y^2 g_q)\, dS_u.$$

We now let $n \to \infty$ to get the first equation. We get the second equation by $d[X^\phi]_t = \phi_y^2 dS_t$ □

The next proposition will serve as an important tool in our analysis.

Proposition 3.2. *Let U, g, τ be as in Proposition 3.1 and let $\forall y, q \in U$*

$$\frac{g_{yy}}{2\phi_y^2} + g_q = 0. \tag{62}$$

Then we can replicate the payoff $g(S_{T\wedge\tau}, [X^\phi]_{T\wedge\tau})$ with the following self financing strategy. We need to hold at each time $t \leq T \wedge \tau$:

$g_y(S_t, [X^\phi]_t)$ shares

$g(S_t,[X^\phi]_t) - S_t g_y(S_t,[X^\phi]_t)$ bonds
The time zero value is $V_0 = g(S_0, 0)$.

Proof. We have for the portfolio value at time $t \le \tau \wedge T$

$$V_t := g(S_t,[X^\phi]_t) - S_t g_y(S_t,[X^\phi]_t) + g_y(S_t,[X^\phi]_t)S_t = g(S_t,[X^\phi]_t).$$

By (3.2) and (62) we have:

$$g(S_t,[X^\phi]_t) = g(S_0,0) + \int_0^t g_y(S_u,[X^\phi]_u)\,dS_u + 0,$$

from this we get the self-financing condition

$$dV_t = g_y(S_t,[X^\phi]_t)dS_t.$$

That the strategy has the claimed time zero value is obvious. □

The condition (62) implies that the payoff $g(S_{T\wedge\tau},[X^\phi]_{T\wedge\tau})$ is just a static bond position and a dynamic stock position. By our proposition, such payoffs can be replicated with the above dynamic stock position.

Now we use this proposition to subreplicate a claim on the stock price received at a time when the log variance X_T of the stock price reaches a barrier.

Proposition 3.3. *Let the stopping time τ_Q be defined as*

$$\tau_Q := \inf(t \ge 0 : [X]_t = Q).$$

For $v \ge 0$ and $y > 0$ and a continuous $f : \mathbb{R}_+ \to \mathbb{R}$ with $|f(e^z)| \le F(e^{|z|})$ for some polynomial F let for $v > 0$

$$BS(y,v,f) := \int_{-\infty}^{\infty} f(ye^z)\frac{1}{\sqrt{2\pi}}e^{-\frac{(z+v/2)^2}{2v}}\,dz$$

and for $v = 0$

$$BS(y,v,f) = f(y).$$

Then we can replicate the payoff at time $T \wedge \tau$:

$$f(S_{\tau_Q})\mathbf{1}_{\tau_Q \le T} + BS(S_T, Q-[X_T], f)\mathbf{1}_{\tau_Q > T} \tag{63}$$

by holding at each time $t \le T \wedge \tau_Q$:

$BS_y(S_t, Q-[X_t], f)$ shares

$BS(S_t, Q-[X_t], f) - S_t BS_y(S_t, Q-[X_t], f)$ bonds
The time zero value of our replicating portfolio is given by $BS(S_0, Q, f)$.

Proof. Our process $[X]_t$ is continuous, since S_t is assumed to be continuous. Therefore τ_Q really is a stopping time.
Let us define $g(y,q) := BS(y, Q-q, f)$.
By direct calculation we see that $\frac{1}{2}y^2 g_{yy} + g_y = 0$ is satisfied on $\mathbb{R}_+ \times (-\infty, Q)$, which is (62) in the log case. The function is in $C^{2,1}$ on U and

continuous on $\bar{U}$. We see that as $q \to Q$ the term $\frac{1}{\sqrt{2\pi}}e^{-\frac{(z+v/2)^2}{2v}}$ converges to the Dirac-delta function. Applying Proposition 3.2 yields our result. $\square$

This result immediately gives a lower bound and a subreplicating strategy on a contract paying

$$f(S_{\tau_Q})\mathbf{1}_{\tau_Q \le T} \tag{64}$$

for some terminal time T and a variance barrier Q.

At time zero we have

$$BS(S_0, Q, f) \le \mathbb{E}[f(S_{\tau_Q})\mathbf{1}_{\tau \le T}].$$

At time t we obtain if $[X]_t < Q$ the lower bound (setting $\tilde{Q} = Q - [X]_t$)

$$BS(S_0, \tilde{Q}, f) \le \mathbb{E}[f(S_{\tau_Q})\mathbf{1}_{\tau \le T}|\mathcal{F}_t]$$

and if $[X]_t > Q$ we obviously have $\mathbb{E}[f(S_{\tau_Q})\mathbf{1}_{\tau \le T}|\mathcal{F}_t] = 0$.

Such contracts have been traded, and are therefore of practical importance.

Obviously, this lower bound is robust in the sense that it holds for all continuous models.

Remark We have seen that in our completely general continuous framework surprisingly the Black Scholes formula comes into play. The above result can be understood as "delta hedging in business time". In a sense, we obtain the Black-Scholes model in our general martingale setting. If $[X]_\infty = \infty$ (which is obviously reasonable), we have by Theorem 1.11

$$X_t = W_{[X]_t} - \frac{1}{2}[X]_t \tag{65}$$

where W is a Brownian motion. Since $X_t = \log(S_t/S_0)$ we find

$$S_t = S_0 e^{W_{[X]_t} - [X]_t/2}. \tag{66}$$

Hence with respect to business time $[X]_t$ the stock price S_t is a geometric Brownian motion.

The next theorem gives a subreplicating strategy for the variance call.

Remark Let us give some intuition to this theorem and assume for the moment that the stopping time τ_Q is bounded.

At the moment when $[X]$ hits Q the variance call turns into a forward starting variance swap starting at the time τ_Q, since the quadratic variation is an increasing process and the call therefore stays in the money.

If $\tau_Q < T$ we need to subreplicate this variance swap. We know how to replicate this variance swap by Theorem 1.6. We can remove the assumption that $-\chi(S_{\tau_Q})$ trades with the help of Proposition 3.3 and find a practical subreplicating strategy.

If $\tau_Q \geq T$ ($[X]$ does not hit Q and the variance call pays zero) we have at time T the payoff $\chi(S_T)$ and a claim on $-\chi(S_{\tau_Q})$. We get

$$\mathbb{E}[\chi(S_{\tau_Q})|\mathcal{F}_T] \leq \chi(\mathbb{E}[S_{\tau_Q}|\mathcal{F}_T]) = \chi(S_T)$$

where the inequality follows from Jensen's inequality and the convexity of χ and the equality follows from the optional stopping theorem and the boundedness of τ_Q. We have shown that our portfolio value is negative and hence less than zero which is the payoff of the variance call. Therefore subreplication has succeeded.

Let us make our arguments precise and prove the main result of this section.

Theorem 3.4. *Let $\chi : (0,\infty) \to [0,\infty)$ be a convex function where the second derivative in the distributional sense has a density χ_{yy}. Assume that it satisfies*

$$\chi_{yy}(y) \leq \frac{2}{y^2} \qquad \forall y \in (0,\infty) \tag{67}$$

and

$$\mathbb{E}[\chi(S_T)] < \infty.$$

We let τ_q be the stopping time from Proposition 3.3 and define

$$\Delta_t = \begin{cases} -BS_y(S_t, Q-[X]_t, \chi) & t \leq \tau_q \\ -\chi_{y-}(S_t) & t > \tau_Q. \end{cases}$$

The following self-financing strategy subreplicates the variance call. At each time $t \leq T$ hold:

1 claim on $\chi(S_T)$

Δ_t shares

$-BS(S_0, Q, \chi) + \int_0^t \Delta_y \, dS_u - \Delta_t S_t$ bonds
The time zero value of the strategy is $V_0 = -BS(S_0, Q, \chi) + \mathbb{E}[\chi(S_T)]$.

Proof. The strategy obviously self-financing. We distinguish the two cases $\tau_Q \leq T$ and $\tau_Q > T$.

If $\tau_Q \leq T$ we have for the portfolio at time T

$$V_T = -BS(S_0, Q, \chi) + \int_0^{\tau_Q} \Delta_u \, dS_u + \int_{\tau_Q}^T \Delta_u \, dS_u + \chi(S_T).$$

Since $\tau_Q \leq T$ we have $\tau_Q \wedge T = \tau_Q$ and Proposition 3.3 gives

$$-BS(S_0, Q, \chi) - \int_0^{\tau_Q} BS_y(S_u, Q-[X]_u, \chi) dS_u = -\chi(S_{\tau_Q})$$

and we obtain

$$-BS(S_0, Q, \chi) + \int_0^{\tau_Q} \Delta_u \, dS_u + \int_{\tau_Q}^T \Delta_u \, dS_u + \chi(S_T)$$

$$= -\chi(S_{\tau_Q}) + \int_{\tau_Q}^{T} \Delta_u \, dS_u + \chi(S_T)$$
$$\leq [X]_T - [X]_{\tau_Q} = ([X]_T - Q)^+$$

where we used Theorem 1.6 with $\phi(S) = \log(S/S_0)$ and $\omega = 1$.
If $\tau_q \geq T$ we have by Proposition 3.3:

$$\begin{aligned} V_T &= -BS(S_0, Q, \chi) + \int_0^T \Delta_u \, dS_u + \chi(S_T) \\ &= -BS(S_T, Q - [X]_T, \chi) + BS(S_T, 0, \chi) \\ &\leq 0 = ([X]_T - Q)^+ \end{aligned}$$

where we used that for a convex function χ BS is increasing in the volatility. To see this calculate:

$$y^2 BS_{yy}(y, q, T) = \frac{1}{qT} BS_q(y, q, T)$$

and differentiation shows

$$BS_y(y, q, T) = \int_{-\infty}^{\infty} \chi_{y-}(ye^z) e^z \Phi'(z, -\frac{1}{2}q^2 T, q\sqrt{T}) \, dz$$

where $\Phi'(y, \mu, p)$ is the normal density with mean μ and standard derivation p. Since χ is convex we get $BS_{yy} \geq 0$ which lead together with the above formula to

$$BS_q(y, q, T) \geq 0.$$

It is obvious that the strategy has the claimed time zero value

$$V_0 = \mathbb{E}[\chi(S_T)] - BS(S_0, Q, \chi).$$

□

The above theorem has shown how to subreplicate the variance call. Since the strategy works for every χ, which satisfies our assumptions we have a whole class of subreplicating strategies. By our theorem we know the time zero value of these strategies. We now seek to find the χ which maximizes this value, and gives a meaningful lower bound for the variance call price in terms of Europeans. The following proposition will tackle this problem. We will see that a certain corridor variance swap is the strategy χ that we are searching for.

Definition 3.5. *Let $K \in \mathbb{R}_+$ and let us define the Black Scholes implied volatility $IV_0(K, T)$ if $K \geq S_0$ as the unique solution to*

$$BS(S_0, IV_0(K, T), (S_T - K)^+) = \mathbb{E}[(S_T - K)^+] \tag{68}$$

and if $K < S_0$ as the unique solution to

$$BS(S_0, IV_0(K, T), (K - S_T)^+) = \mathbb{E}[(K - S_T)^+]. \tag{69}$$

Proposition 3.6. *Let $\chi : (0,\infty) \to [0,\infty)$ be a convex function where the second derivative in the distributional sense has a density χ_{yy}. Assume that it satisfies*

$$\chi_{yy}(y) \le \frac{2}{y^2}. \qquad \forall y \in (0,\infty). \tag{70}$$

Furthermore, we assume that the function $K \mapsto IV_0(K,T)$ is Borel measurable. We then have:

$$\mathbb{E}[\chi(S_T)] - BS(S_0,Q,\chi) \le \mathbb{E}[\chi^{opt}(S_T)] - BS(S_0,Q,\chi^{opt}) \tag{71}$$

where

$$\chi^{opt}(y) = -2\log(y/S_0)\mathbf{1}_{\{y\in\mathbb{R}^+ : IV_0(y,T)>Q\}}. \tag{72}$$

Proof. Note that by setting $F = S_0$ in Theorem 1.12 we get by setting $C_K(y) = (y-K)^+$ and $P_K(y) = (K-y)^+$:

$$\mathbb{E}[\chi(S_T)] = \chi(S_0) + \int_0^{S_0} \mathbb{E}[P_K(S_T)]\chi_{yy}(K)\,dK + \int_{S_0}^{\infty} \chi_{yy}(K)\mathbb{E}[C_K(S_T)]\,dK$$

where the term with the first derivative cancels because S_t is a martingale and $C(S_0) - P(S_0) = S_T - S_0$. By Theorem 1.12 we also have:

$$\begin{aligned} BS(S_0,Q,\chi) = \chi(S_0) &+ \int_0^{S_0} \chi_{yy}(K) BS(S_0,Q,P_K(S_T))\,dK \\ &+ \int_{S_0}^{\infty} \chi_{yy}(K) BS(S_0,Q,C_K(S_T))\,dK. \end{aligned}$$

Using these results we get:

$$\begin{aligned} \mathbb{E}[\chi(S_T)] - BS(S_0,Q,\chi) &= \int_0^{S_0} \chi_{yy}(K)[\mathbb{E}[P_K(S_T)] - BS(S_0,Q,P_K(S_T))] \\ &\quad + \int_{S_0}^{\infty} \chi_{yy}(K)[\mathbb{E}[C_K(S_T)] - BS(S_0,Q,C_K(S_T))] \\ = \int_0^{S_0} &\chi_{yy}(K)[BS(S_0,IV_0(K,T),P_K(S_T)) - BS(S_0,Q,P_K(S_T))]\,dK \\ + \int_{S_0}^{\infty} &\chi_{yy}(K)[BS(S_0,IV_0(K,T),C_K(S_T)) - BS(S_0,Q,C_K(S_T))]\,dK \end{aligned}$$

where we have used the definition of the implied volatility for the second equation.

We have that $0 \le y^2\chi_{yy}(y) \le 2$ by the convexity of χ and (70).
Since the Black-Scholes formula is increasing in volatility the term in the brackets is positive if and only if $IV_0(K,T) > Q$. Therefore, we have

$$\int_0^{S_0} \chi_{yy}(K)[BS(S_0,IV_0(K,T),P_K(S_T)) - BS(S_0,Q,P_K(S_T))]\,dK$$

$$+ \int_{S_0}^{\infty} \chi_{yy}(K)[BS(S_0, IV_0(K,T), C_K(S_T)) - BS(S_0, Q, C_K(S_T))]\, dK$$
$$\leq \int_{\{K \geq S_0 : IV_0(K,T) > Q\}} \frac{2}{K^2}\mathbb{E}[C_K(S_T)]\, dK$$
$$+ \int_{\{K < S_0 : IV_0(K,T) > Q\}} \frac{2}{K^2}\mathbb{E}[P_K(S_T)]\, dK.$$

The last expression makes sense since the integrals are finite by our assumption $\mathbb{E}[[X]_T] < \infty$ and the assumption that IV_0 is measurable. We obviously have by (28):

$$\mathbb{E}[\chi^{opt}] = \int_{\{K \geq S_0 : IV_0(K,T) > Q\}} \frac{2}{K^2}\mathbb{E}[C_K(S_T)]\, dK + \int_{\{K < S_0 : IV_0(K,T) > Q\}} \frac{2}{K^2}\mathbb{E}[P_K(S_T)]\, dK$$

which completes the proof. □

Remark For example a continuous function with countably many jumps is still measurable, thus the assumption that IV_0 is measurable has little practical impact.

As for the previous volatility derivatives, we can get information on their price from a continuum of European vanilla options. The way to subreplicate the variance call reminds us of the formula for the variance swap. We need to hold the same European option positions for a strike K if we have $IV_0(K,T) < Q$, if not we do not hold options with this strike. Given all available option prices of the market we can model independently find a lower bound for the price of the variance call by approximating:

$$\int_{\{K \geq S_0 : IV_0(K,T) > Q\}} \frac{2}{K^2} C(K)\, dK + \int_{\{K < S_0 : IV_0(K,T) > Q\}} \frac{2}{K^2} P(K)\, dK.$$

Note that as for the variance swap case, the above strategy is practical in the way that we only have to continuously trade in the stock and hold a static option position. However, in contrast to the variance swap case, we merely find a subreplication strategy instead of an exact replication strategy under these general assumptions.

Remark The above results hold at time zero. What about a general time t? If we want to find a lower bound for the price of a variance call with strike Q and expiry date T at some time $t \in (0,T)$, we distinguish between two cases.
If $[X]_t > Q$ we will surely finish in the money (since $[X]_t$ is increasing), so we can exactly price the variance call using the strategy of the variance swap since we have the payoff $[X]_T - Q$.
If $[X]_t \leq Q$ we use the above method with the the new strike $\tilde{Q} = Q - [X]_t$.

A trivial upper bound for the variance call price is the variance swap price given by $-2\mathbb{E}[\log(S_T)/S_0)]$. For more information on upper bounds see Carr and Lee [7].

3.3. **A log-normal assumption.** After having tried to tackle the problem of pricing the variance call model-independently, we make a very strong assumption.
In this section we assume that the for some time $t \leq T$ the remaining volatility $\sqrt{[X]_T} - \sqrt{[X]_t}$ is log-normal distributed, i.e. $\log(\sqrt{[X]_T} - \sqrt{[X]_t})$ is $N(\mu_t, s_t^2)$ distributed. Empirical studies show that this assumption is indeed very reasonable.

We set:

$[X]_{0,t}$: the running variance from time 0 to t

$VOL_t := \mathbb{E}[\sqrt{[X]_T}|\mathcal{F}_t]$ the time t price of the volatility swap

$VAR_t := \mathbb{E}[[X]_T|\mathcal{F}_t]$ the time t price of the variance swap

Our assumption implies:

$$\mathbb{E}[\sqrt{[X_T]} - \sqrt{[X]_t}|\mathcal{F}_t] = e^{\mu_t + s_t^2/2}$$

where μ_t and s_t^2 denote the time-t conditional parameters of our log-normal distribution.

From the above equation we get:

$$\mathbb{E}[(\sqrt{[X_T]} - \sqrt{[X]_t})^2|\mathcal{F}_t] = e^{2\mu_t + 2s_t^2}.$$

We then have:

$$\mu_t + s_t^2 = \log(VOL_t - \sqrt{[X]_{0,t}})$$

and

$$2\mu_t + 2s_t^2 = \log(VAR_t - 2VOL_t\sqrt{[X]_{0,t}} + [X]_{0,t})$$

which we solve for μ_t and s_t to obtain:

$$s_t^2 = \log(VAR_t - 2Vol_t\sqrt{[X]_{0,t}} + [X]_{0,t}) - 2\log(VOL_t - \sqrt{[X]_{0,t}})$$

and

$$\mu_t = 2\log(VOL_t - \sqrt{[X]_{0,t}}) - 2\log(VAR_t - 2Vol_t\sqrt{[X]_{0,t}} + [X]_{0,t}).$$

We see that we can compute the time t-conditional mean and variance of our log-normal distribution from the prices of variance and volatility swaps. These prices can be calculated with methods from the preceding sections or obtained from the market.

We then get from Carr and Lee [9] the following Black-Scholes style formula for the price of the variance call:

(73)

$$\mathbb{E}[([X]_T - K)^+|\mathcal{F}_t] = \begin{cases} VAL_t - K & K \leq [X]_{0,t} \\ v_2 N(d_0) + 2\sqrt{[X]_{0,t}} v_1 N(d_1) - (K - [X]_{0,t}) N(d_2) & K > [X]_{0,t} \end{cases}$$

where:

$$v_1 := VOL_t - \sqrt{[X]_{0,t}}$$

$$v_2 = VAR_t - 2VOL_t\sqrt{[X]_{0,t}} + [X]_{0,t}$$

and

$$d_j := \frac{m_t - \log(\sqrt{K} - \sqrt{[X]_{0,t}})}{s_t} + (2-j)s_t. \quad j = 0, 1, 2$$

The great advantage of this formula is that all that is required to compute the variance call price are the current volatility swap and variance swap prices and the observable running variance at time t. Then a simple calculator can do the calculation without problems. The drawback, of course, is the strong assumption of the log-normal distribution.

We will numerically test the robustness of this formula by underlying a specific model. We will use the Bates model of the previous section. The following tables gives an overview of the parameters we have used for our tests.

Figure	λ	α	δ	θ	ξ	κ	T	strike
11	0.1	-0.1;-0,3-0,5	0.2	0.04	0.39	1.05	2	0 to 0.2
12	0.1	-0.1 to -0.7	0.2	0.04	0.39	1.05	2	at the money
13	0.1	-0.1 to -0.7	0.2	0.04	0.39	1.05	2	75% of swap
14	0.1	-0.1 to -0.7	0.2	0.04	0.39	1.05	2	40% of swap

In Figure 11 we plot the true value and (73) against the strike for three different values of α. Obviously, increasing the jump size increases the value of the variance call. We see that for very low strikes we have almost no error (if the strike is zero we have a variance swap and the error is zero). Then we have some strikes where (73) overestimates the true value and finally for very high strikes (73) underestimates the true value. We find this pattern for all jump sizes for roughly the same strikes.

In Figure 12 we see the important at the money variance call (the call where $K = VAR_0$) plotted against the jump size α. We see that for small jump sizes the at the money variance call is still in the area of strikes where (73) slightly overestimates the true value and for higher jump sizes we have an underestimation of the true value. Note that Figure 11 also suggests this behavior. We have the same plot in Figure 13 with $K = 0.75\,VAR_0$ and we see that the errors are significant.

Figure 11 suggests that for calls deep in the money formula (73) should be a very good approximation. By Figure 14 we see that for the strike $K = 0.4\,VAR_0$ we have nearly no error for reasonable small jump sizes ($\alpha \leq 0.4$).

Remark In summary, we have seen that for deep in the money strikes formula (73) is a very good approximation of the true price of the variance call when the price follows Bates dynamics. Additionally, we have found that how deep the strike has to be in the money depends on the jump size.

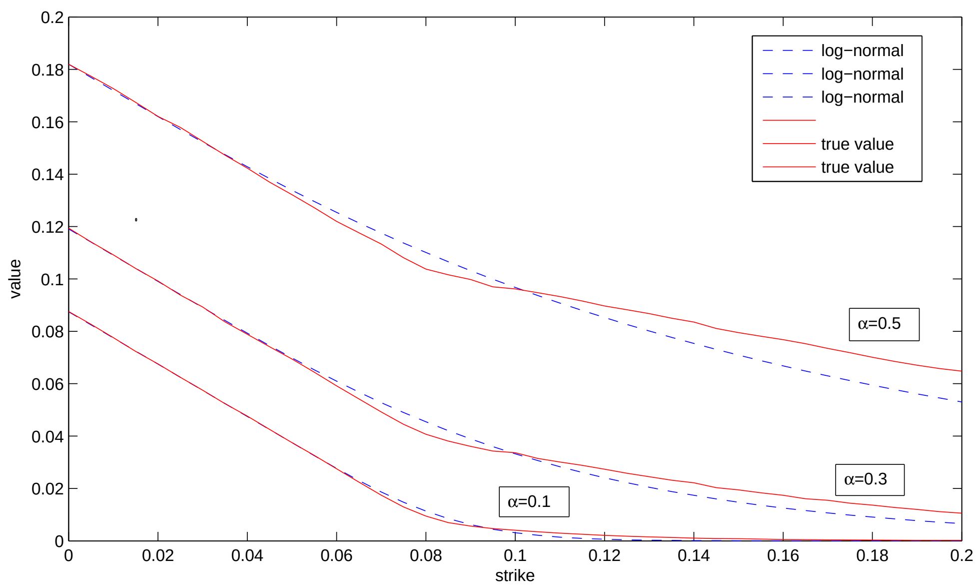

FIGURE 11. Bates model-plotted against the strike K

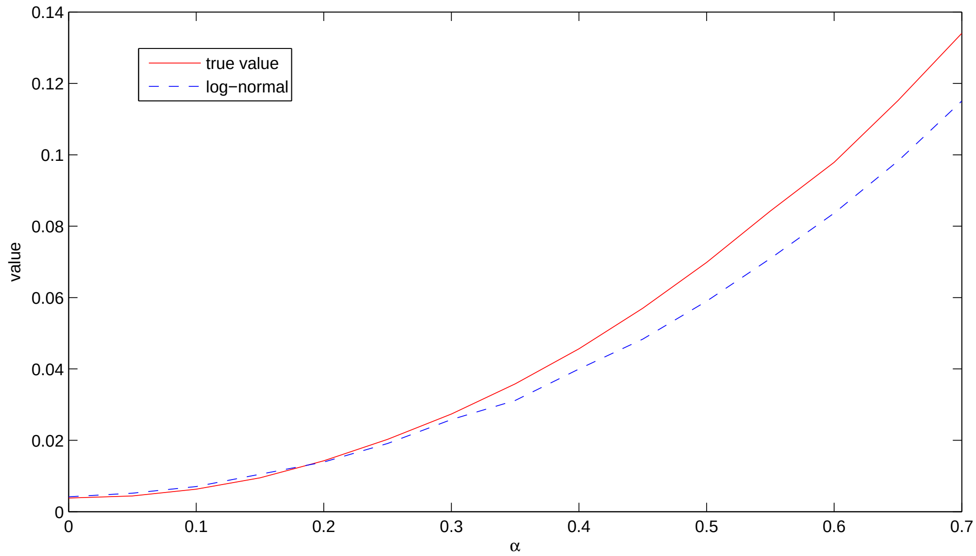

FIGURE 12. at the money variance call plotted against jump size α

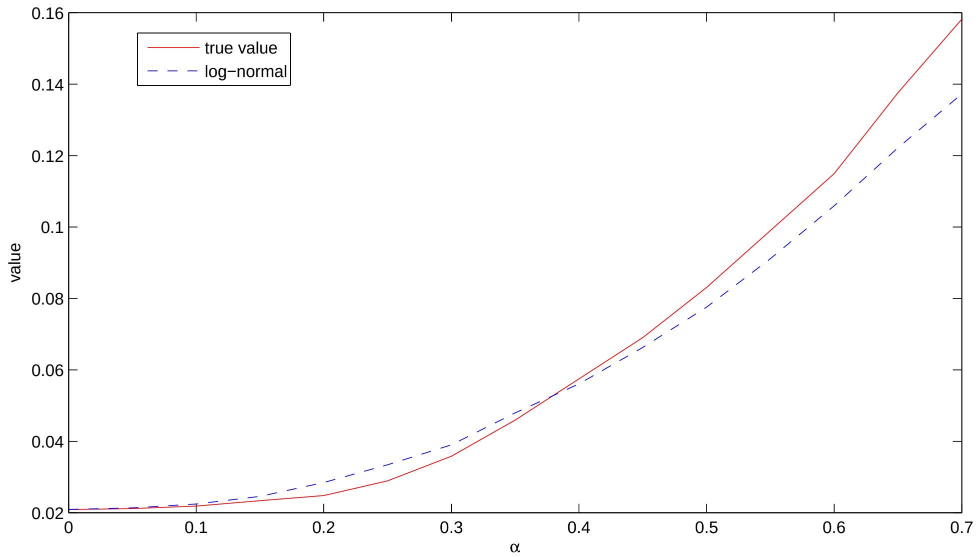

FIGURE 13. in the money variance call plotted against jump size α

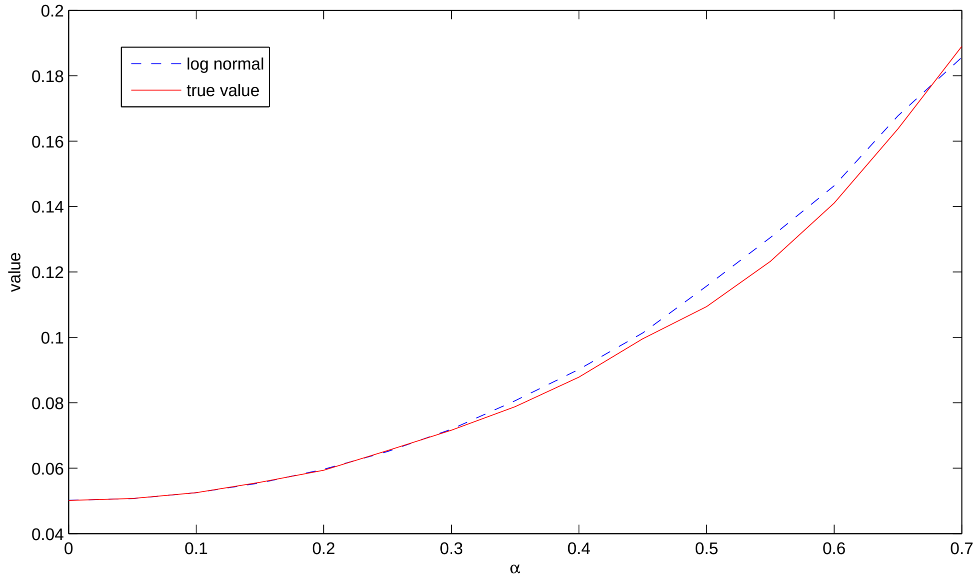

FIGURE 14. in the money variance call plotted against jump size α

References

[1] Leif Andersen and Jesper Andreasen *Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing, Review of Derivatives Research 4, 231-262, 2000*

[2] Gurdpi Bakshi, Charles Cao, and Zhiwu Chen *Empirical performance of alternative option pricing models, Journal of Finance, 52, 2003-2049, 1997*

[3] Andrew Matytsin *Modeling volatility and volatility derivatives, Download from www.math.columbia.edu/smirnov/Matytsin.pdf 1999*

[4] Peter Carr, Eric Chang, Dilip Madan *The Variance Gamma Process and Option Pricing,European Finance Review 2: 79105, 1998*

[5] Peter Carr and Dilip Madan *Towards a theory of volatility trading, in R.Jarrow, editor, Volatility, pages 417-427, Risk Publications, 1998*

[6] Peter Carr and Roger Lee *Robust Replication of Volatility Derivatives, Bloomberg LP and University of Chicago, Download from http://math.uchicago.edu/ rl/rrvd.pdf, 2008*

[7] Peter Carr and Roger Lee *Hedging Variance Options on Continuous Semimartingales, Finance and Stochastics, vol 14 issue 2, 179-207, 2010*

[8] Peter Carr, Roger Lee, and Liuren Wu *Pricing Variance Swaps on Time-Changed Lévy Processes, Finance and Stochastics , Forthcoming, 2009*

[9] Peter Carr and Roger Lee *Realized Volatility and Variance: Options via Swaps, RISK, vol 20 issue 5, 76-83, 2007*

[10] Peter Carr and Roger Lee *Put-call symmetry: Extensions and applications, 2008, Mathematical Finance, Vol. 19, No. 4, 523560, 2009*

[11] Darrel Duffie, Jun Pan, and Kenneth Singleton *Transform analysis and asset pricing for affine jump diffusions, Econometrica 68, 1343-1376, 2000*

[12] Filo Fiorani *Option Pricing under the Variance Gamma Process , Merrill Lynch & Co., 2004*

[13] Jim Gatheral *The Volatility Surface, John Wiley & Sons, Inc.m Hoboken, New Jersey, first edition, 2006*

[14] Steven L. Heston *A closed form solution for options with stochastic volatility, with application to bond and currency options, Review of Financial Studies 6, 327-343, 1993*

[15] Franz Hofbauer *Höhere Wahrscheinlichkeitstheorie, lecture notes*

[16] Andrey Itkin *Pricing options with VG model using FFT, Cornell University Library, arXiv:physics/0503137v2, 2005*

[17] Daniel Revuz and Marc Yor *Continuous Martingales and Brownian motion ,Sringer-Verlag, Berlin Heidelberg New York, third edition, 2005*

[18] Tina Hviid Rydberg*The normal inverse gaussian lévy process: simulation and approximation, Stochastic Models,13:4,887-910, 1997*

[19] Wim Schoutens *Lévy processes in finance, John Wiley & Sons, Inc.m Hoboken, New Jersey, first edition, 2003*

Curriculum Vitae

Personal data

Name:	Michael Punz
Date of Birth:	25^{th} of September, 1987
Place of Birth:	Baden, Lower Austria
Citizenship:	Austria

Education

03/2007-01/2010:	Diploma studies in Mathematics, University of Vienna Specialisation: Stochastics and Applications
06/2006:	Matura - Final school examination
1998 to 2006:	Bundesrealgymnasium Berndorf, high school

Work Experience

01/2011-02/2011:	PricewaterhouseCoopers, Vienna Consulting, Financial Services Group
09/2010-12/2010:	Financial Market Authority Austria, Vienna Actuarial Issues and Models Group
08/2009:	Volksbank, Baden Back office, IT-Organisation
08/2008:	Volksbank, Baden Back office, IT-Organisation